

RESULT 13

US-09-702-705-784
 ; Sequence 784, Application US/09702705
 ; Patent No. 654010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; Fanger, Gary
 ; Lodes, Michael A.
 ; APPENDANT: Bangur, Chaitanya S.
 ; APPENDANT: Bangur, Chaitanya S.
 ; APPENDANT: Fanger, Gary
 ; APPENDANT: Lodes, Michael A.
 ; APPENDANT: Vedula, Tom
 ; APPENDANT: Carter, Darrick
 ; APPENDANT: Retter, Marc
 ; APPENDANT: Mannion, Jane
 ; APPENDANT: Fan, Liqun
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 784
 ; LENGTH: 6353
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-702-705-784

Query Match Score 106.6; DB 4; Length 6353;
 Best Local Similarity 50.5%; Pred. No. 7e-21;
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 103 GCTGTACCACTACTGCCCTGAAGAACCCGTTACCGTACATGGACTGTGCCGTACTACCTG 162
 Db 5324 GCGTCAAATGGCAATTAGAAAGCCGGTTACCATATGATTCTGCACATGGTACAT 5383

Qy 163 AACGAGGGTAGGTGGTGGAGGTATCCGGTACATGGAGGAAACCCCTCGGTGAAG 222
 Db 5384 AATGAGGAGCAGGTGGACTGCCATCGGAGTGGAGATGCCAGTGTAAG 5440

Qy 223 CGTGAGGACATCTCTGACCAAGGTGGAGCACCTCCACCGTTATGAGGAGCTC 282
 Db 5441 AGAGAGAGACADATTCPACCTTCAGCTTCAAGCTTGGACCAATTCCATGCC 5500

Qy 283 CTCTGGTCATGGACACTCTGAGGCTTCTGAGGCTTGGACTGTACGTGATAATGTCCTC 342
 Db 5501 CGACCAAGCTTGGAGAAGGTACTGAAAATCTCAATTGGACTATGTTGACCTATCTT 5560

Qy 343 GTTCACTGGCCATTCTGGCGAGAGAAATGCCAGGGTGAAGCTGGCCCTGAC 402
 Db 5561 ATTCAATTCTCAAGTGTCTGTAAGCAGGTAGGAG-----TGTATCCAAA 5608

Qy 403 GGAAATAACGTCAATTCTCAGGACCTGACCCACATGGCTTCAACTGGACATGCC 522
 Db 5609 GATGAAATGAAAAATACATTACAGTGTGGATCTGTGACATGGAGGCCATG 5668

Qy 463 GAGANGATTATGAGGATCAGCAAGGCCAGGTCATGGCTTCAACTGGACATGCC 522
 Db 5669 GAGAAGCTGAAAGATCAGGATGGCAACTCCATGGGTGCAACTTCAACAGG 5728

Qy 523 GACCTCTGGAGAGATGTCAGCTTCGCTTCAACGCCACAGATC 576
 Db 5729 CTGCTGGAGATGATGATCTCAACAGCCGGCTCAAGTACAAGCTGTGCAACCAGGTG 5788

Qy 577 GAGATTCAACCCCTCTGGCCAACAGGAGCTGGTCACTGTTCTCAAGAACATT 636
 Db 5789 GAATGTCATCTTACTTCAACCAGAAACTGTCGGATTCGAAAGTCAAAGACATT 5848

RESULT 14

US-09-736-457-784
 ; Sequence 784, Application US/09736457
 ; Patent No. 650448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPENDANT: Bangur, Chaitanya S.
 ; APPENDANT: Fanger, Gary
 ; APPENDANT: Lodes, Michael A.
 ; APPENDANT: Vedula, Tom
 ; APPENDANT: Carter, Darrick
 ; APPENDANT: Retter, Marc
 ; APPENDANT: Mannion, Jane
 ; APPENDANT: Fan, Liqun
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C15
 ; CURRENT APPLICATION NUMBER: US/09/736,457
 ; CURRENT FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 1864
 ; SEQ ID NO: 784
 ; LENGTH: 6353
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-736-457-784

Query Match Score 106.6; DB 4; Length 6353;
 Best Local Similarity 50.5%; Pred. No. 7e-21;
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 103 GCTGTACCACTACTGCCCTGAAGAACCCGTTACCGTACATGGACTGTGCCGTACTACCTG 162
 Db 5324 GCGTCAAATGGCAATTAGAAAGCCGGTTACCATATGATTCTGCACATGGTACAT 5383

Qy 163 AACGAGGGTAGGTGGTGGAGGTATCCGGTACATGGAGGAAACCCCTCGGTGAAG 222
 Db 5384 AATGAGGAGCAGGTGGACTGCCATCGGAGTGGAGATGCCAGTGATGGTACAT 5440

Qy 223 CGTGAGGACATCTCTGACCAAGGTGGAGCACCTCCACCGTTATGAGGAGCTC 282
 Db 5441 AGAGAGAGACADATTCPACCTTCAGCTTCAAGCTTGGACCAATTCCATGCC 5500

Qy 283 CTCTGGTCATGGACACTCTGAGGCTTCTGAGGCTTGGACTGTACGTGATAATGTCCTC 342
 Db 5501 CGACCAAGCTTGGAGAAGGTACTGAAAATCTCAATTGGACTATGTTGACCTATCTT 5560

Qy 343 GTTCACTGGCCATTCTGGCGAGAGAAATGCCAGGGTGAAGCTGGCCCTGAC 402
 Db 5561 ATTCAATTCTCAAGTGTCTGTAAGCAGGTAGGAG-----TGTATCCAAA 5608

Qy 403 GGAAATAACGTCAATTCTCAGGACCTGACCCACATGGCTTCAACTGGACATGCC 522
 Db 5609 GATGAAATGAAAAATACATTACAGTGTGGATCTGTGACATGGAGGCCATG 5668

Qy 463 GAGANGATTATGAGGATCAGCAAGGCCAGGTCATGGCTTCAACTGGACATGCC 522
 Db 5669 GAGAAGCTGAAAGATCAGGATGGCAACTCCATGGGTGCAACTTCAACAGG 5728

Qy 523 GACCTCTGGAGAGATGTCAGCTTCGCTTCAACGCCACAGATC 576
 Db 5729 CTGCTGGAGATGATGATCTCAACAGCCGGCTCAAGTACAAGCTGTGCAACCAGGTG 5788

Qy 577 GAGATTCAACCCCTCTGGCCAACAGGAGCTGGTCACTGTTCTCAAGAACATT 636
 Db 5789 GAATGTCATCTTACTTCAACCAGAAACTGTCGGATTCGAAAGTCAAAGACATT 5848

Db	500	AATGAGGAAGCAGTTCAGTTGACTCGCTCCACCGTTAGGAGACCTG 556	Query Match Similarity 10.9%; Best Local Similarity 50.5%; Pred. No. 5.5e-21; Indels 30; Gaps 4; Matches 386; Conservative 0; Mismatches 349;
Qy	223	CGTGAGGACAACTTCGTGCAACAGGTGGAAACCCTCCACCGTTAGGAGACCTG 282	Qy 103 GCTGTCAACAGTCGCCCTGAAAGCCGGTACCGTCACTGGACTGCTGCTACTACCTG 162
Db	557	AGAGAGAACATTCATTCACATTCACCTCAAGCTTGGACCAATTCCCATCGACAGACTGGTC 616	Db 440 GCGGTCAAAATGGCAATAGAGAGCCGGTTCCACCATTTGATTCAGATGTTAACAT 499
Qy	283	CTCTGGTCATGACCACTCTGAAAGCTTGAACCTAAGTGTGATATGGTCTCTC 342	Qy 163 AACGAGGGTAGGGTGAAGTTGGGGTATCGTSACTTCTGTGAGGAGAACCCCTGGTAAG 222
Db	617	CGACCAAGCTTGGAAAGGTGACTGAAAATCTCAATTGGACTATGGTGCCTCTATCT 676	Db 500 AATGAGGAGGAGCTTGGACTGGCCCATCG 556
Qy	343	GTTCACTGGCCCATGGCCAGGAGAAATGGCCAGGTGGCCATGAC 402	Qy 223 CCGTGGAGACATCTTCGTCACCAAGGTGGTGAACCAACCTCCACCGTTAGGAGCTG 282
Db	677	ATTCAATTTCAGTGTGCTGTGTTGGCTTGGCAAGTGGAGAAGTGGAGAAGTGGAGA 724	Qy 557 AGGAAAGACATATTCTACACTCTAAAGCTTGGAGAATTCCATGACCGAGTGTGGTC 616
Qy	403	GGCAAATACTGTCATTCGAAGGACCTGACCAGAAACCCGAGGCCACATCGCGCTATG 462	Db 283 CTCGGTCCCATGACGACTCCCTGAAAGCTTGGACTCTGACTATGTGTCCTC 342
Db	725	GATGAAAATGGAAAATACTTTGACAGCTTGGTGCACATGGAGGCCATG 784	Qy 617 CGACCAAGCTTGGEEAGGTCACTGAAATCTCAATTGGACTATGTGCTCATCTT 676
Qy	463	GAGAAGATTATGAGGATTCAGGCAAGCCATGTCCTCCACATTGCC 522	Qy 343 GTTCACTGGCCCATGGTGCACAGTAAGGAAATGGCTTGGCCZAGATGGCCCTGAC 402
Db	785	GAGAAAGTGAATGAGATTGGCAAGTGGCATGGGTTGTCAACTCAACAGCG 844	Db 677 ATTCAATTTCAGTGTGCTGTGTTGGAGGAAAGCTGGTGGAGGAG 724
Qy	523	GACCTTGAGAGATGTCCAAGTTGCTCAGGCTTCAACGAGTC 516	Qy 403 GGAAATAATGCTGATTTCTCAAGGACCTTCAAGGACTTCAACCTGAGGAAACCCGAGCCCACATGGCCTATG 462
Db	845	CTGCTGGAGAGATCTTCAGAAGCCAGGGTCAGTACAGGCTTCTCAACAGGTG 904	Db 725 GATGAAAATGGAAAATACTATTGACAGCTTGGGATCTGGCACTATGGGAGGCCATG 784
Qy	577	GAGATTCACCCCTTCCTGCCAACAGGAGGAGTGGTGCAGACTGTGTTCTCAAGAACATT 636	Qy 463 GAGAAGGATTATGAGGATGTCAGGCAAGGCCAGGTCACTGGACCATTTGCC 522
Db	905	GAATGTCATCTTACTTCACCGAGAAACCTGTTCTGGAAGTGAAGAACATT 964	Db 785 GAGAAGTGTAAAGATGGAGGATGGGAGATGGTCACTGGGTGTCACACTCAACCAAGG 844
Qy	637	ATGCCCGTGGCTACTCTCTGGGCTTCGGAGAACCAAGGTTCC-----GACCAAC 687	Qy 523 GACCTTGAGAAGATGTCACAGTTCAGGAACTGGCTTCAGGAGACTGCTCTCCAAAGAACATT 576
Db	965	GTTCTGTGTTGCTATAGTGTCTGGATCCATCAGAAACATGGTGGACCGAAC 1024	Db 845 CTGCTGGAGAGATCTTCCTCAAGGCCATCTCAAGGCAAGGGCTCAAGGAGTC 904
Qy	688	GCTGAGCGGGTAGGGAGAACAGACTCTGAAAGGATCCGGAGAAAGCCGGAAACCC 747	Db 577 GAGATTTCACCCCTTCCTGCCCAACAGGAGGTGGTCACTGCTCTCCAAAGAACATT 636
Db	1025	TCCCCGGTGTGCTTGGAGGCCAGTCGCTTGTGCTTGGAAAAAAAGCAACGGAAACC 1084	Db 905 GAATGTCATCTTACTTCACCAAGGAGAAAATGGCTTCTGCAAGTCAAAGAACATT 964
Qy	748	CTTGCTCAGGTTCTTATTGCTGGGTTCTGGCTTACCGCTGTTCTCCCAAGAGGC 807	Db 637 ATGCCCGGGCTACTCTCTGCTTCTGCAAGAACCCAGTTCC-----CAGCACC 687
Db	1085	CCAGCCCTGTGTTGCCCTGGCTTACCAAGCTGAGCCGGGGTTGTCCTGGCAAGGCC 1144	Db 965 GTCCTGGTGCCTATAGTGTCTGGATTCGGATCAAGGAGAACCATGGTGGACCGAAC 1024
Qy	808	TCAAACCCCAAGGCATGGTCAACTTAAGGAGTTAGCTC 852	Qy 688 GGTGAGGGGGTCAAGGAGAACAGACTCTGACGAGATGGGGGCAACACC 747
Db	1145	TACAATGAGCAGGCGCATAGACAGAACGGTGTGTTGAAATTG 1189	Db 1025 TCCCGGTGCTCTGGGACCCAGTCCTTGTGCTTGGAAAAAAAGCAACGGAAC 1084
RESULT 10			Qy 748 CTGGCTCAGGTTCTTATTGCTGGGTTCTGGCTTACGGTGTCTCCCAAGAGGC 807
US-09-614-124B-807			Db 1085 CCAGCCCTGTGTTGCCCTGGCTTACAGCTGAGCCTGGTCTGGCAAGAGGC 1144
Sequence 807, Application US/09614124B			Qy 808 TCCAACCCCAAGGCATGGTCAACTTAAGGAGTTAGCTC 852
Patent No. 6630574			Db 1145 TACAATGAGCAGGCGCATAGACAGAACGGTGTGTTGAAATTG 1189
GENERAL INFORMATION:			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
FILE REFERENCE: 21012.47BC9			
CURRENT APPLICATION NUMBER: US/09/614,124B			
NUMBER OF SEQ ID NOS: 1668			
SEQUENCE ID NO: 807			
LENGTH: 3629			
ORGANISM: Homo sapiens			
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND			
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER .			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Fan, Linjun			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Fan, Linjun			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger, Gary			
APPLICANT: Vedwick, Tom			
APPLICANT: Carter, Darrick			
APPLICANT: Rettler, Marc			
APPLICANT: Mannion, Jane			
APPLICANT: Bangur, Chaitanya S.			
APPLICANT: Lodes, Michael A.			
APPLICANT: Fanger,			

Qy 523 GACCTTGAAAGATGTCAGTCCAACTGTCG-----CAAGTCATGCCCTACGCCAACAGATC 576
 Db 796 CTGCTGAGATGATCTCAAAAGCCAGGGTCAAGCTGTGCAATCGGC 855
 Qy 577 GAGATTCAACCCTTCCTGCCAACAGGAGCTGCTGCAAGAACAT 636
 Db 856 GATGTCACTCTTAACGAAAACCTGTGGATTCTGAAGTAAAGACAT 915
 Qy 637 ATGCCCTGGCTTAACCTCTGGCTGGCTGGAGAACAGETTC-----OACCCAC 687
 Db 916 GTCTGTGCTTACTAGCTGGATCCCACGAAAGAACCTGGTGGACAC 975
 Qy 688 GGTGAGGGGTAGCGAGAACAGAAGACTGACAGATCCCGAAGGGCGAACAC 747
 Db 976 TCCCCGGTGMCTTGGAGAACCCAGTCCTTGCTGCAAGGAAAC 1035
 Qy 748 CTTGCTAGGTCTTATTGCCCTGGGCTGGCTACGTGTTCCCAAAGAGC 807
 Db 1036 CCAAGCCCTGATGCCCTGGCTTACCGCTGAGCTGGGTCTGGCCAAGAGC 1095
 Qy 808 TCCAACCCAAAGCGATTGASTCCAACTCAAGAGCATGAGCTC 852
 Db 1096 TACAATAGCGAGGCATAGCAGAACGTCAGGTGTTGAAATIC 1140

RESULT 8
 US-09-702-705-807
 Sequence 807, Application US/09702705
 Patent No. 6504010
 GENERAL INFORMATION:
 APPLICANT: Wang, Tongtong
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodes, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedick, Tom
 APPLICANT: Carter, Darrick
 APPLICANT: Rettner, Marc
 APPLICANT: Mannion, Jane
 APPLICANT: Fan, Liqun
 APPLICANT: Pan, Liqun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121_478C14
 CURRENT APPLICATION NUMBER: US/09/702,705
 CURRENT FILING DATE: 2000-10-30
 NUMBER OF SEQ ID NOS: 1833
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 807
 LENGTH: 3829
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-702-705-807

Query Match Score 10.9%; Score 106.5; DB 4; Length 3829;
 Best Local Similarity 50.5%; Pred. No. 5.5e-21; File Reference: 210121_478C14
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 103 GCTGTACCACTGCCCCTGAAGACCGGGTACCGTACTTGGTACCTCTG 162
 Db 440 GCGTCAATTGCAATAGGAGGTATCGTGTACCTTGGTACCTCTG 499
 Qy 163 AACGAGGGTGAAGCTGGTGAAGGAGCTTCTGAAGGAGAACCCCTCGGTGAAG 222
 Db 500 AATGAGGAGCAGGTGGACTGCCATCG--AAGAAGATTGCAATGGCAAG 556

Query Match Score 10.9%; Score 106.6; DB 4; Length 3829;
 Best Local Similarity 50.5%; Pred. No. 5.5e-21; File Reference: 210121_478C14
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 223 CGTGGAGGACATCTCTGCAACCAAGGTGGAACCACTCCACGGTATGAGGACGTC 282
 Db 557 AGRAAGACATATTCTACACTCAAACTTGGGAAATTCCCATGACCAAGATGGTC 616
 Qy 283 CTTGGTACATGAGCAGTCCCTGAAGGCTTGGACTAGTTGATATGTTCTC 342
 Db 617 CGACCGGCTTGTGAAAAGGTCACTGAAATACTTCATTTGGACTATTTGACCTCTATCT 676
 Qy 343 GTCACTGGCCATTGTCGCGAGAACATGGCAGGATGGCCCTGAC 402
 Qy 677 ATTCACTTCCAGTGTAAAGCCAGGTGGAG-----TGATCCAAAAA 724
 Qy 403 GGCAGAATAGTCATTCTCAAGGACCTGACGGAAACCCGAGCCACATGGCGGTATG 462
 Db 725 GATGAAATGGAAAATAACTATTGACACAGTGGATCTCTGGCACTGGGACATGGGCATG 784
 Qy 463 GAGAGATTATGAGATGCAAGGGAGGTCACTGGTGTCTCAACTGSACCATGCC 522
 Db 765 GAGAGTAAAGTAGTGGAGTTGGCAAGTCATGGGTCTCAACTTCACZCAGG 844
 Qy 523 GACCTGAGAAAGATGTCAAAGTGCCTGCGCAACGAGATC 576
 Db 845 CTGGGGAGATGCTAACAGCAGGCTCAAGTACAGCTAAAGCCTCTGAAACAGGTG 904
 Db 577 GAGATTCACCCCTTCCTGCCAACGGAGACTGGTCACTGCTGAGTCTCTCAAGACATT 636
 Db 905 GATGTGATCCCTTACTTCACCAAGAGAAAATCTGCTGATTTGCAAGTCAAAGCATT 964
 Qy 637 ATGCCGTGGCTACTCTCTGGTGTGCAAGAACAGGTTC-----CACACC 687
 Db 965 GTCCTGCTGGCTATGCTCTGGATCCATGGTGAACCTGGTCTCCCAAGAGC 807
 Db 1085 CCAGCCCTGATGCCCTGGCTTACCGCTGAGCTGGGGGGCAACACC 747
 Qy 688 GGTTGAGGGTCAKGAGAACAGACTCTGAAAGGAGATGGGGGGTGTGCTGCTGGCAAGAGC 1024
 Db 1145 TCCCCGGTGGCTCTGGAGGCCAGCTCTTGTGCCTTGGAAAAAAGCAAGGAAAC 1084
 Db 1025 TCCATGGAGCTAGCACTGAGGACCTGAGCTGGTGTGCTTGTGCCTTGGAAAAAAGCAAGGAAAC 1084

RESULT 9
 US-09-736-457-807
 Sequence 807, Application US/09736457
 ; General Information:
 ; Patent No. 6509448
 ; General Information:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Rettner, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Ajun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121_478C15
 ; CURRENT APPLICATION NUMBER: US/09/736,457
 ; CURRENT FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 1864
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 807
 ; LENGTH: 3829
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-736-457-807

Query Match Score 10.9%; Score 106.6; DB 4; Length 3829;
 Best Local Similarity 50.5%; Pred. No. 5.5e-21; File Reference: 210121_478C15
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 103 GCTGTACCACTGCCCCTGAAGACCGGGTACCGTACTTGGTACCTCTG 162
 Db 440 GCGTCAATTGCAATAGGAGGTATCGTGTACCTTGGTACCTCTG 499
 Qy 163 AACGAGGGTGAAGCTGGTGAAGGAGCTTCTGAAGGAGAACCCCTCGGTGAAG 222
 Db 500 AATGAGGAGCAGGTGGACTGCCATCG--AAGAAGATTGCAATGGCAAG 556

Query Match Score 10.9%; Score 106.6; DB 4; Length 3829;
 Best Local Similarity 50.5%; Pred. No. 5.5e-21; File Reference: 210121_478C15
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 103 GCTGTACCACTGCCCCTGAAGACCGGGTACCGTACTTGGTACCTCTG 162
 Db 440 GCGTCAATTGCAATAGGAGGTATCGTGTACCTTGGTACCTCTG 499
 Qy 283 CTTGGTACATGAGCAGTCCCTGAAGGCTTGGACTAGTTGATATGTTCTC 342
 Db 617 CGACCGGCTTGTGAAAAGGTCACTGAAATACTTCATTTGGACTATTTGACCTCTATCT 676
 Qy 343 GTCACTGGCCATTGTCGCGAGAACATGGCAGGATGGCCCTGAC 402
 Qy 677 ATTCACTTCCAGTGTAAAGCCAGGTGGAG-----TGATCCAAAAA 724

RESULT 6
 US-09-542-615A-171
 / Sequence 171, Application US/09542615A
 / Patent No. 6518256
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Tongtong
 / APPLICANT: Fan, Lijun
 / APPLICANT: Kalos, Michael D.
 / APPLICANT: Bangur, Chaitanya S.
 / APPLICANT: Hosken, Nancy A.
 / TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 / CURRENT APPLICATION NUMBER: US/09/542,615A
 / CURRENT FILING DATE: 2000-04-14
 / NUMBER OF SEQ ID NOS: 350
 / SOFTWARE: FastSEQ for Windows Version 3.0
 / SEQ ID NO: 171
 / LENGTH: 1491
 / TYPE: DNA
 / ORGANISM: Homo sapien
 US-09-542-615A-171

Query Match Score	10.9%	Score 106.6;	DB 4;	Length 1491;
Best Local Similarity	50.5% ;	Pred. No. 3.5e-21;		
Matches	386; Conservative	0; Mismatches 349;	Indels 30; Gaps 4;	

Qy 103 GCTGTACCACTGCCTGAAGACCGGTATTACGGTCACTGGACTGTGCCGTACTACCTG 162
 Db 391 GCGGTCRAATTCGGCAATAGAAGCCGGTRCCACCATATTCCTGCACATTTACAT 450

Qy 163 AACGAGGTGAGGTTGAGGTATCCGTGACTCTCTGAGGAGAACCCCTCGGTGAG 222
 Db 451 AATGAGGAGGAGGTTGAGCTGGCCATCCTG--AACAAAGTTGAGATGCACTGTGAG 507

Qy 223 CGTGAGGACAACTTCGTCGACCAAGGTGAGGTATGGAAACCCTCACCGTTATGAGGACCTC 282
 Db 508 AGAGAAAGACATATTCTGACATTCAGCTTGGACAATTCCATGCAAGACTTGTCTC 567

Qy 283 CTCTGGTCATTGACCACTCCTGAAAGGCTCTGGACTGACTACGTGATAATGTTCTC 342
 Db 568 CGACCAAGCTGGAAAGGTCACTGAAAATCTCAATTGGACTATGTTGACCTCTATCTT 627

Qy 343 GTTCACGGCCATTGCTGGAGGAGAATGGCCAGGGTAGGCCAAGATGGCCCTGAC 402
 Db 628 ATTCAATTCTCAGTGTCTGTAAGCCAGGTGGAGA-----TGTACCCAAA 675

Qy 403 GCCAATAATAGCTCATTCGAAGGACCTGACCGGAAACCCGAGCCACATSGGCCGTATG 462
 Db 676 GATGAAATGAAATAATCATTTGACATGGTCTTGTGCACTGGGGCCATG 735

Qy 463 GAGAAGATTATGAGGATCGCAAGGCGAGGTGCACTGGTCTCCAACTGGACATTGCC 522
 Db 736 GAGAATGTAAGATCAGGATGGCAAGGTCATGGGTCAACTCAACACAGG 795

Qy 523 GACCTTGAGAAGATGTCGACATTGCG---CAGGTCTGACATGGTCTTGTGCACTGGGGCATTG 576
 Db 796 CTGCTGGAGATGATCCTCAAAAGCCAGGGCTCAAGTCAAGCTGCAACAGGTG 855

Qy 577 GAGATTCAACCCCTCTGCCCCAGGAGGACTGTGGTCACTGTCTCAGAACACTT 636
 Db 856 GAATGTCATCCTTACTTCACACAGAAACTGTCGTTGARTTCTGAAAGTCAAAAGACATT 915

Qy 637 ATGCCCTGGCTACTCTCTGGCTCGGAGAACCGGTTCC-----CACCAC 687
 Db 916 GAGAATGTTAAAGATGGAGATGGGATCTGGATCCACGAGAAACCATGGGTGAC 975

RESULT 7
 US-09-606-421B-171
 / Sequence 171, Application US/09606421B
 / Patent No. 6531315
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Tongtong
 / APPLICANT: Fan, Lijun
 / APPLICANT: Kalos, Michael D.
 / APPLICANT: Bangur, Chaitanya S.
 / APPLICANT: Hosken, Nancy
 / APPLICANT: Panger, Gary R.
 / APPLICANT: Li, Samuel X.
 / APPLICANT: Wang, Ajun
 / APPLICANT: Skeiky, Yasir A. W.
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 / TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 / FILE REFERENCE: 210121.455C9
 / CURRENT APPLICATION NUMBER: US/09/605,421B
 / CURRENT FILING DATE: 2000-06-28
 / NUMBER OF SEQ ID NOS: 358
 / SOFTWARE: FastSEQ for Windows Version 3.0
 / SEQ ID NO: 171
 / LENGTH: 1491
 / TYPE: DNA
 / ORGANISM: Homo sapien
 US-09-606-421B-171

Query Match Score	10.9%;	Score 106.6;	DB 4;	Length 1491;
Best Local Similarity	50.5% ;	Pred. No. 3.5e-21;		
Matches	386; Conservative	0; Mismatches 349;	Indels 30; Gaps 4;	

Qy 103 GCTGTACCACTGCCTGAAGACCGGTATTACGGTCACTGGACTGTGCCGTACTACCTG 162
 Db 391 GCGGTCRAATTCGGCAATAGAAGCCGGTRCCACCATATTCCTGCACATTTACAT 450

Qy 163 AACGAGGTGAGGTTGAGGTATCCGTGACTCTCTGAGGAGAACCCCTCGGTGAG 222
 Db 451 AATGAGGAGGAGGTTGAGCTGGCCATCCTG--AACAAAGTTGAGATGCACTGTGAG 507

Qy 223 CGTGAGGACAACTTCGTCGACCAAGGTGAGGTATGGAAACCCTCACCGTTATGAGGACCTC 282
 Db 508 AGAGAAAGACATATTCTGACATTCAGCTTGGACAATTCCATGCAAGACTTGTCTC 567

Qy 283 CTCTGGTCATTGACCACTCCTGAAAGGCTCTGGACTGACTACGTGATAATGTTCTC 342
 Db 568 CGACCAAGCTGGAAAGGTCACTGAAAATCTCAATTGGACTATGTTGACCTCTATCTT 627

Qy 343 GTTCACGGCCATTGCTGGAGGAGAATGGCCAGGGTAGGCCAAGATGGCCCTGAC 402
 Db 628 ATTCAATTCTCAGTGTCTGTAAGCCAGGTGGAGA-----TGTACCCAAA 675

Qy 403 GCCAATAATAGCTCATTCGAAGGACCTGACCGGAAACCCGAGCCACATSGGCCGTATG 462
 Db 676 GATGAAATGAAATAATCATTTGACATGGTCTTGTGCACTGGGGCCATG 735

Qy 463 GAGAAGATTATGAGGATCGCAAGGCGAGGTGCACTGGTCTCCAACTGGACATTGCC 522
 Db 736 GAGAATGTAAGATCAGGATGGCAAGGTCATGGGTCAACTCAACACAGG 795

Qy 523 GACCTTGAGAAGATGTCGACATTGCG---CAGGTCTGACATGGTCTTGTGCACTGGGGCATTG 576
 Db 796 CTGCTGGAGATGATCCTCAAAAGCCAGGGCTCAAGTCAAGCTGCAACAGGTG 855

Qy 577 GAGATTCAACCCCTCTGCCCCAGGAGGACTGTGGTCACTGTCTCAGAACACTT 636
 Db 856 GAATGTCATCCTTACTTCACACAGAAACTGTCGTTGARTTCTGAAAGTCAAAAGACATT 915

Qy 637 ATGCCCTGGCTACTCTCTGGCTCGGAGAACCGGTTCC-----CACCAC 687
 Db 736 GAGAATGTTAAAGATGGAGATGGGATCTGGATCCACGAGAAACCATGGGTGAC 975

APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121_455C11
 CURRENT APPLICATION NUMBER: US/09/643,597
 CURRENT FILING DATE: 2000-08-21
 NUMBER OF SEQ ID NOS: 369
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 171
 LENGTH: 1491
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-643-597-171

Query Match 10.9%; Score 106.6; DB 4; Length 1491;
 Best Local Similarity 50.5%; Pred. No. 3.5e-21;
 Matches 386; Conservative 0; Mismatches 349; Indels 30; Gaps 4;

Qy 103 GCTGTCAACACTGCCCTGAAACCGTTACCGTCACTGGACTGCGCTGGTACTACCTG 162
 Db 391 GCGTCACAATTGCAAATAAGCCGGTTCCACCATATTGATTGCGCACATGTTAACAT 450
 Qy 163 AACAGGGTGAGGTGGTGGAGGGTATCGTGACTCTGGAGGAGAACCCCTCGGTGAAG 222
 Db 451 AATGAGGAGCAGGTGGACCTGGCATGGCCATCG--AAGGAAGATGCGATGCGTGAAG 507
 Qy 223 CCTGAGGACATCTTCGCTGCACCAAGGTGGAAACCACCTCCACCGTATGAGGACGCTG 282
 Db 508 AGAGAAACATATTCTACACITCAAGCTTGGAAATTCATGACCAAGTGGTGGTC 567
 Qy 283 CTCCTGGTCACTTGACGACCTCCCTGAAAGGCTCTGGACTGACTAGTTGATATGTTCC 342
 Db 568 CGAACAGCCTGGTGAAGGTCACTGAAAAAATCTCATTTGAGCTATGTTGACCTCTPATCT 627
 Qy 343 GTTCACTGGCCATTGTCGCCAGAGAATGCCAGGTGGCCAGGTGGCCCTGAC 402
 Db 628 ATTCAATTCCAGTGTCTGTAAGGCCAGTGGAGA-----TGATCCAAAA 675
 Qy 403 GCGAAATAAGTGTATTCTCAAGGACCTGAGGAACTGGCCACAGCCGCTATG 462
 Db 676 GATGAAATGGAAAAAAATCATTTGACACAGTGGATCTGTCATGGAGGCCATG 735
 Qy 463 GAGAAAGATTATGAGGATGCGAAAGGCCAGGTCACTGGTCTCAACTGGACCATGGC 522
 Db 736 GAGAGTGTAAAGTGGAGGGTTGGCAAGTGTGGGTGTCACACTCAACACAGG 795
 Qy 523 GACCTTGAGAAGATGTCCAAGTTCGCAAGGCAACAGATC 576
 Db 796 CTGCTGAGAATGTCCTCAACAGCCAGGGTCAAGTACAGGCCGTGTCACAGGTG 855
 Qy 577 GAGATTGACCCCTTCTGCCAACGGAGACTGGTCCAGACTGCTTCTCAAGAACAT 636
 Db 856 GAATGTCATCTTACTCAACAGAGAAACTGTCAGTCAAAGAACATT 915
 Qy 637 ATGCCCTGGCTTACTCTCTGGGTCTGGTCAAGAACAGGTTC-----CAGCACC 687
 Db 916 GTTCTGCTGTTGCTTATAGTCGCTGGATCCCATCGAGAACCCATGGTGGACAC 975
 Qy 688 GTTGTGGGGTCAAGGAGAACAGACTCGAACGAGATGGCGGACACCCAC 747
 Db 976 TCCCCGGTGCCTTGGAGGACCCAGTCCTTGTGCTTGGCAAAAGCAAGGAAAC 1035
 Qy 748 CTTGCTGTTCTTATTGCTGGGCTGCCCCTGGCTACGTCCTTCTCCCAAGAGC 807
 Db 1036 CCAGCCCTGATGTCGCCCTGGTCAACAGCTGAGTGGGTGTCCTGGCAGAGC 1095
 Qy 808 TCCAACCCAAAGGCCATTGASTCCTAACTCAAGAGGATTAAGCTC 852
 Db 1096 TACAATGAGGGCATAGAGAACGTCGAGGCTTGTGAGGACCCAGTCCTTGTG 1140
 Qy 976 TCCCCGGTGCCTTGTGAGGACCCAGTCCTTGTGAGGACCCAGTCCTTGTG 1035
 Db 1036 CCAGCCCTGATGCCCTGGTCAACAGCTGAGTGGGTGTCCTGGCAGAGC 1095

Qy	504	CTCCAATGAGGACCATTCGCACACTTGAGAAGATGTCAA-----GTTGCCAAGGTAT	557	Matches 424 ; Conservative 0 ; Mismatches 380 ; Indels 33 ; Gaps 5 ;
Db	512	CTCCAACTTCAACCCTCTCTGAGATGGGATCTTGAAACAACCTGGTTAACGTATA	571	Qy 24 ATTGAGCAACCGCGTCAAGATTCCTCGTCGCRTRGGTACCTTGGTCAAGGTTC 83
Qy	558	GCCCTACGCCAACAGATCGGATTCACCCCCTCTGCCAACGAGGACTGGTCCAGNA	617	Db 53 ACTCAACAGGACCAAATGSCCCACCTGGTCACTGGACTCT 106
Db	572	GCTGCTGTTAACGATCGGATCTGAGTCACTTAATCAGAGAACCTGATTGAGA	631	Qy 84 CAAAGGGAGACCTATCTGCTTCAACCCTGCGCTAACGCGGTACCTCTGGAA 143
Qy	618	CTGCTCTCCAAAAGAACATATGCCCGTGGCTACTCTCTGGGCTCGAGAACCGT	677	Db 107 TCTGCGGAGGTGACCGAGCTGAGGGTCTATGACATGGGATTCGCACTTGA 166
Db	632	CTGCCATTGCAAGGATCGGGTGAATCTGACATCCCTGTTCTGACAGGCC	691	Qy 144 CTGTGCTGGTACTCTGCTTCAAGAGGTGAGGTTGGTCACTGGTACCTCTGGAA 203
Qy	678	TCCCACCCACCGCTGAG-----CGGTCAACGAGAACACTTGAGAGATGCCA	731	Db 167 CTGGCCCCAGGTTGACAGATGAGAGGCTGGGTGCGCTTCAGGAGAACCTCA 226
Db	692	CTGGGCAAGGCTGGGACCCCTCTCTGGAGGATTCGGAGAACGAAATGAC	751	Qy 204 GGAAACCCCTGGTGAAGCAGGACATCTGGTTCGAGCAACAAAGTGTGAAACCACCT 263
Qy	732	GAAGGGGGGACACCCCTGTCAGTCACTGGTCTATGGTCAAGTGA	791	Db 227 GGAGCAGT-----GGTAGAGGCGCAGATCTCTCATGTCAGCAAGCTGTTGACGT 283
Db	752	CAAGTACATAAAACTACAGGCCAACCCCAGGCCATTAGTCAACTCAAGAGCATTA	811	Qy 264 CCACCGTTGAGGACATCTGGTCACTGGCCATTATGGTGGCCAGAAAGATGGCGAGTGA 323
Qy	792	C GTTCTCCCAAGAGCTCACAAGCCATTAGTCAACTCAAGAGCATTA	848	Db 284 CCGAGACAGAGATGTAAGGGCTGCGAGAGCTGAGGACCTGAGCTGGAA 343
Db	812	C GTGATCCCCAAGCTGTCGACACGACCACTATGTCAGAACCTTAAAGTTGA	868	Qy 324 CTACGTGATACTGTTCTCGTCACTGGCCATTCTGGTGGCCAGGGTGA 383
Db	344	CTACCTGGACCTTACCTTACTGCACTGGCAACTCTGGCTTCTGGCTTCTGG 391		Qy 384 GCCAAAGTTGGCCCTACGGAAATAATAGCTTCAAGGACCTGATTCTCAAGGAAACCCGA 443
Qy				Db 392 GCCTGACTATTTCCTGGCATGGGAAACCTGATTCCTGACACCGATTGGT 451
Db				Qy 444 GGCACATGGGGCTATGGAAAGATTAGGGTCGAAGGGCTGCAATTGGTGT 503
Db				Db 452 GGACACTGGGGCTTGGCAACTAGTGTGAGGTTGGTAAAGGAATGGAGT 511
Qy				Qy 504 CTCACACTGGACCATTCGGACCTTGGAAAGATGTCAA-----GTTGCCAAAGGTCA 557
Db				Db 512 CTCACACTTCACCCCTTCAATTGAGGATCTGACAAACCTGGCTTAAGATAA 571
Qy				Qy 558 GCCTCACCCAAAGATGAGATTACCCCTCTGCACTCTGGCTGGAGATA 617
Db				Db 572 GCCTGCTTTAACAGATGCACTGACTGCAACCCATACCTAAC 631
Qy				Qy 618 CGCTCTCCAGAACATTGCGCTTACTCTCCCTGCGTCAAGGAGT 677
Db				Db 632 CTGCCATTGCAAGGATCGGGTGAATCTGACATACCTCCCTGGTCTCTGACAGGCC 691
Qy				Qy 678 TCCACCCRCGCTGAG-----CGGTCAAGGAGAACAGACTCTAACGAGATCCCGA 731
Db				Db 692 CTGGGCCAGGCTGAGCCCTGAGGATTCGGGATTAAGGAATTCAGC 751
Qy				Qy 732 GAAGGGCGCAACCCCTGGTCACTGGCTGGCTGGCTACGT 791
Db				Db 752 CAGTACATAAAACTCAAGCCAGGCTGCACTGGTCAACGGAAACTGGT 811
Qy				Qy 792 CCTCTCCCCAAGAGCTCAACCCAAAGCGATTGAGTCCAAACTTAAGAGCATGA 848
Db				Db 812 CGTGTGACACCCAACTCTGGTCAAGGACGTTAAGGTAAAGCTTGA 868
RESULT 4				
US-09-498-599-3				
Sequence 3, Application US/09498599				
; GENERAL INFORMATION:				
; APPLICANT: Cameron, Douglas C.				
; ADDRESS: DeWitt Ross & Stevens S. C.				
; STREET: 8000 Excelsior Drive, Suite 401				
; CITY: Madison				
; STATE: WI				
; COUNTRY: U.S.A.				
; ZIP: 53717-1914				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: Patent-in Release #1.0, Version #1.30				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/09/498,599				
; FILING DATE:				
; CLASSIFICATION:				
; NAME: Sara, Charles S.				
; REFERENCE/DOCKET NUMBER: 09820-037				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 608-831-2100				
; TELEFAX: 608-831-2106				
; INFORMATION FOR SEQ ID NO: 3:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 1337 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: DNA (genomic)				
; HYPOTHETICAL: NO				
; ANTI-SENSE: NO				
; ORIGINAL SOURCE: Rat lens aldose reductase gene				
; ORGANISM: Rat				
; Query Match Score 113; DB 4; Length 1337;				
; Best Local Similarity 50.7%; Pred. No. 4.5e-23;				
; Sequence 171, Application US/09643597				
; GENERAL INFORMATION:				
; APPLICANT: Bangur, Chaitanya S.				
; ADDRESS: 6426072				
; APPLICANT: Hosken, Nancy R.				
; APPLICANT: Li, Samuel X.				
; APPLICANT: Wang, Ajun				
; APPLICANT: Skeky, Yasir A.W.				

TITLE OF INVENTION: FROM SUGAR
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dewitt Ross & Stevens S.C.
 STREET: 8000 Eexcelior Drive, Suite 401
 CITY: Madison
 STATE: WI
 COUNTRY: U.S.A.
 ZIP: 53717-1914
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/801,344
 FILING DATE: 08/08/801,344
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Sara, Charles S.
 REFERENCE/DOCKET NUMBER: 09820.037
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 608-831-2100
 TELEFAX: 608-831-2106
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1327 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Rat lens aldose reductase gene
 US - 08-801-344-3

Query	Match	Start	End	Score	DB	Length
Query	Match	11-6%	Score 113;	DB 3;	Le	
Best Local Similarity	Similarity	50.7%	Pred. No. 4	5e-33;		
Matches	Matches	424	Conservative	0;	Mismatches	380;
YY	YY	24	ATTGACCAAGGGCGTCAAGATTCCTGGCTCGGCTTGTTGGTGT			
YY	YY	53	ACTCAACAGGACCAAGATGCCACCTGGGTCTGGCAAC			
YY	YY	84	CAAGGGCGAGACCTATAATGCTGTCACCATGCCCTGAAGG			
YY	YY	107	TCCGGCCCAAGTGAACCGCAGGCTGTGAAGGTCTCTATGCAC			
YY	YY	144	CTGGGCGCTGTAACCTGACAGGGTGAACGGGCTGAGCTTGTTGG			
YY	YY	167	CTGGGCCCAAGTGTACCAAGATGAGAAGGTGGGGTGG			
YY	YY	204	GGAGAACCCCTCGCTGAGCTGAGGACATCTCGTGTGCA			
YY	YY	227	GGAGGAGT--GGTAGGGCCAGGATCTCTTCTATGTC			
YY	YY	264	CCACCGTTATGAGGACAGTCTCTGGCTCATGAGACTCTCC			
YY	YY	284	CCACGACCAAGGATGGTAAAGGGCTGGAGAAGACGGGTT			
YY	YY	324	CTACGGTGTATGTTCCCTGGATGCTGGCCATTGTCGGCC			
YY	YY	344	CTACCTGGACCTACCTTACTGGCAACTG-----			
YY	YY	384	GCCCCAGATGGCCCTGAGGGCAATACGTCATTCTGAAGG			
YY	YY	392	GCCGACTATTCCTGGATGCTGGAAACGGTATTGTC			
YY	YY	444	GCCCCATGGCCCTATGGAGAGATTATAGGATGCG			
YY	YY	452	GGACACTTGGACGCGCTATGGACAACTGGATGAGGTT			

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2004, 01:05:39 ; Search time 76 Seconds
 (without alignments)
 7141.342 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978

Sequence: 1 aetcctaaggaaagactt.....ccaaggacccgtctggatgaa 978

Scoring table: IDENTITY NUC
 Gapov 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : Issued Patents NA:
 1: /cgns_6/picodata/2/ina/5A-COMB.seq/*
 2: /cgns_6/picodata/2/ina/5B-COMB.seq/*
 3: /cgns_6/picodata/2/ina/6A-COMB.seq/*
 4: /cgns_6/picodata/2/ina/6B-COMB.seq/*
 5: /cgns_6/picodata/2/ina/PCITS-COMB.seq/*
 6: /cgns_6/picodata/2/ina/backfile1.seq/*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULTS

RESULT 1
 US-09-023-655-1010
 ; Sequence 1010, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocke, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seelhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023, 655
 ; FILING DATE: HEREDITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1010:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1335 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: S179035
 ; US-09-023-655-1010
 ; Query Match 12.2%; Score 119.2%; DB 4; Length 1335;

Result No.	Score	Query Match	Length	DB ID	Description
1	119.2	12.2	1335	4 US-09-023-655-1010	Sequence 1010, App
2	113	11.6	1337	3 US-09-001-344-3	Sequence 807, App
3	113	11.6	1337	4 US-09-098-559-3	Sequence 3, App
4	106.6	10.9	1491	4 US-09-643-597-171	Sequence 171, App
5	106.6	10.9	1491	4 US-09-180-884A-171	Sequence 171, App
6	106.6	10.9	1491	4 US-09-042-615A-171	Sequence 171, App
7	106.6	10.9	1491	4 US-09-606-421B-171	Sequence 171, App
8	106.6	10.9	3829	4 US-09-702-705-807	Sequence 807, App
9	106.6	10.9	3829	4 US-09-736-447-807	Sequence 807, App
10	106.6	10.9	3829	4 US-09-614-14B-807	Sequence 807, App
11	106.6	10.9	3829	4 US-09-345-884-807	Sequence 807, App
12	106.6	10.9	3829	4 US-09-089-14A-807	Sequence 807, App
13	106.6	10.9	6353	4 US-09-702-745-784	Sequence 784, App
14	106.6	10.9	6353	4 US-09-736-447-784	Sequence 784, App
15	106.6	10.9	6353	4 US-09-614-14B-784	Sequence 784, App
16	106.6	10.9	6353	4 US-09-071-345-784	Sequence 784, App
17	106.6	10.9	6253	4 US-09-589-14A-784	Sequence 784, App
18	103.8	10.6	1265	4 US-09-702-705-95	Sequence 95, App
19	103.8	10.6	1265	4 US-09-736-445-95	Sequence 95, App
20	103.8	10.6	1265	4 US-09-514-14B-95	Sequence 95, App
21	103.8	10.6	1265	4 US-09-671-345-95	Sequence 95, App
22	103.8	10.6	1265	4 US-09-14A-95	Sequence 95, App
23	103	10.5	1199	4 US-09-319-112-80	Sequence 80, App
24	103	10.5	1199	4 US-09-976-54-1019	Sequence 1079, App
25	101	10.1	1316	3 US-09-147-803-11	Sequence 323, App
26	99.2	10.1	1316	4 US-09-702-705-323	Sequence 323, App
27	99.2	10.1	1316	4 US-09-124B-323	Sequence 323, App

QY 227 AGGACATCTCGTGTGACCAAGGTGTGAAACCACCTCCACGGTTATGAGCACGTCCTCT 286
 Db 297 AGGAATGTCATCRAAACAAGCTGTGAACTCAGACCC-----ATGGCATGTGCTTG 350
 QY 287 GGTCCATTGAGCACTCCCTGAAAGGTCTTGACTGTGACTAGTTGATGTTACCTCGCTTC 346
 Db 351 AAGCTGCAGGACGCCCTGAAAGCTCAGTAACTCTGACCTCTAACCTCATCC 410
 QY 347 ACTGCCCATTGTCGCGAGAAGATGGCAGGTGAGCCAGATAAGCTGCGCA 406
 Db 411 ATTCCCGATGAGCACTCCACACCGGAGTCGCAAGACTCTAGGCCTCGGTGACG 470
 QY 407 AATAAGTCATTCAGGACCTGAGCAAGCTGAGGCAACATGGCGCTATGGAA 466
 Db 471 ATGGGTGCTGGAATCGAACCATATCTCTGGAAACGATGGCAGGGATGAAAG 530
 QY 467 AGATTATGAGGATGCAAGCCAGGTCCATTGGTGTCTCACTGGGACCC 526
 Db 531 AGCTGTTTCATGGGTGTGTCAGTGAATGCAACTACACATCTCTCA 590
 QY 527 TTGAAAGATGTCCTGAAAGTGGCTCAAGGCACTGCTGCTGAACTTGCACC 586
 Db 591 CCAGAGACTCTGCTGCAATAAGGGCTAGAACATGGGAGACGACACC 650
 QY 587 CCTTCTGCCAACCGAGGAGCTGTGACTACTGC--TTCTCAAGAACATTATGCCG 643
 Db 651 CCTACTTCCGGCCACTCTTCCTCAACTTCTCCAGAGCAAGGGATCTGGTGACCC 710
 QY 644 TGGCTACTCTCCMCTGGGTCCACACCGGTTGGGGTAGGGTAGGGCG 703
 Db 711 CGCACACCCCCGCTGGTGGTCTCACCGCCAACCGCAACGGGTGGGAGCTCGTGC 770
 QY 704 AGAACAGACTCTGAACTGAA---TGCCTGGAGGGCGCAACCCCTGGTCAGGTT 760
 Db 771 TGGAGACCTGTATCAAGTGTGGTGTATAGTGGAGACGGGCGCAGCTCG 830
 QY 761 TTATGCCCCTGGGTCTGCCGTGCTACTGCTGTTCTCCAGAGCTCAACCCAAAGC 820
 Db 831 TCCCTCGCTGGGACGGGACAACGGTGGGATECCAGACCTCAAGGTGAGA 890
 QY 821 GCATGAGTCAACTTCAGGACATTGAGCTCTCCGATGCGACTTGA 869
 Db 891 GGCTGAGGAACTTGAAGTGTGCTGTTGACATCTCGEGGAGGA 939

QY 19 TTACATGACCAACGCCGTCAGATTCCAGGTTGGTACCTTGCTGCTAGTCAA 78
 Db 2 TTGCTCTCAACCCGGCAAGTCCTGCGGCTGGGACTCTGGCTGGCTGGCT 55
 QY 79 GGTCCAGGGAGACCTAATCTGTTACACTGCTGCTGAAACCGTTACCGTC 138
 Db 56 GCGGAAACGGCTGCTGGGACGCCGTTATGCGGCTGAAAGCACTTACGGAC 115
 QY 139 TTGACTGTGCTGTTGACTACTGTGAGGTTGGAGGTATCGTCACTTC 198
 Db 116 ATTGACTGTGTCAGGATCAAACAAAGGAAAGGGTGGGTGTCAGGAGTC 175
 QY 199 CTGAAGGAGAACCCCTGGTAAGCTGAGGACATCTCTGCAAGGTGTGAAAC 258
 Db 176 CTGGAGGGCATA--GTCAGGGTGAAGATCTCTGCACTCAGCTTGGGAC 232
 QY 259 CACCTCACCCTTAAAGGACGTCCTGCAATTGAGCACTCCCTGAAAGCTTGGAA 318
 Db 233 ATAAACATGGGGGAGGAGTCCGGTGTGCACTGAAGCACTGGAGACTTGCA 292
 QY 319 CTTGACTACGGTTGATATGTTCTCTGCTGGCCCATTTGCTGCCGAGAAAGATGGCCAG 378
 Db 293 ACTGACTACCTGGATCTCTACCTCATGCTGGCC----- 327
 QY 379 GGTGAGCCAAAGATGCCCTGACGCCAAATACGTCTCTGACGGAC 438
 Db 328 ----GGTGGGGATGAAAGAAGGGCCGGTTTCGCCCGAACGCGCTCAAC 382
 QY 439 CCCGAGGCCCATGGGGCTATGGAAAGATTATGAGGATGCAAGGGCACTCCAT 498
 Db 383 ATCCCCGGGAGCTGGGGCGCATGGAGGCTGAGGCTGAGGGGGCCCATC 442
 QY 499 GGTGTCCTCAACTGGACATGGGACTCTGCGGACTCTGGCAAGTGTGATG 558
 Db 443 GGGTAGAAGCAAGTCTCTGACCAAGAGCTGGGAGCTGGGGTGGCC 502
 QY 559 CCTCAGGCCAACGAGATGGGATCTGGGACTCTGGCCGCTGCACTAC 618
 Db 503 CCCGGCTGACCTGGGAGCTGGGAGGCTGGGGCTGGGGCTGGGGCT 562
 QY 619 TGCTTCTCCAGAACATATGGCCCTGCTACTCTCTGCTGCTGCTGCTG 678
 Db 563 TGCGCTCCCAAGGATACCTCTCGCCCTACTCGCCGCTGCTGCTGCTG 622
 QY 679 CCCACCACCGGTAGCGAGAAAGAGCTGAGGATTCGGCAAGGAGC 738
 Db 623 ACGGTGAAGGGGGGGGCTGAGAACCCGGCGCTGCTGGGGCTGGCTG 682
 QY 739 GCGAACACCCCTTGCTGCTGGCTCTTATGGCTGGGGCTGGCTGCTTC 798
 Db 683 GGAAGAGCTCCAACCCAGGCTGCTGCTGGCTGCTGGGGCTGGGCAAGATGGGCAACGGTGC 742
 QY 799 CCCAAGAGCTCCAACCCAGGCTGCTGCTGGGGCTGGCTGCTGGGGCTGGCTG 839
 Db 743 CGAAAGAGCCGACGAGGGAGGATAGGGCAATCTCGCTGCTGCTGCTG 783

RESULT 15
 US-10-425-114-30950
 ; Sequence 3059, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yimua
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21153313)B
 ; CURRENT APPLICATION NUMBER: US10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO: 30950
 ; LENGTH: 1185
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZNFBLB73106F07_FLI
 US-10-425-114-30950
 ; Query Match Score 129; DB 13; Length 1185;
 ; Best Local Similarity 50.8%; Pred. No. 1.2e-30;
 ; Matches 417; Conservative 0; Mismatches 365; Indels 39; Gaps 3;

Search completed: May 29, 2004, 13:07:48
 Job time : 499 secs

RESULT 13
; Sequence 17144, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 17144
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3068-021-G2_FLI
US-10-425-114-17144

Query Match Score 130.8; DB 13; Length 1289;
Best Local Similarity 52.4%; Pred. No. 3 4e-31; 0; Mismatches 357; Indels 21; Gaps 5;

Qy 107 TCACCACTGCCTGAAGACCCGTTACCTGACTTGGCTGTGCTGACTCTGTAAC 166
Db 185 TCCACACAGCCTCCGTGCGTACGCCACCTGACTGCGCCCTGACTACAGAAG 244
Qy 167 AGGGTGAGGTTGGTGGGGPATTCCTGACTTCCTGAAGAGAAACCCCTGGTGAAGCTG 226
Db 245 AAGCTGAGGTTGGTGAAGGGATTCAAGACGGACT -- GTCAAGGGCG 301
Qy 227 AGGACATCTTCTGACCAAGGTGTGACCAACTCCACCGTTATGAGGAACCTCCTCT 286
Db 302 AGGACCTGTTCATCACAAACCAAGCTGAGCTGAGCTAGACCTGCTG 355
Qy 287 GTCCATTGAGCTCCGAGGCTGTTGACTGAGTGTGATATGTTCTCGTCTG 346
Db 356 AAGGCTGCAAGAACCCGTTGAGGCTGAGCTGAGCTGAGCTACCTCAAC 415
Qy 347 ACTGGCCATTTGCTGCCGAGAGATGCCGGGTGAGCCAGATGGCCCTGACGGA 406
Db 416 ATTCCTGAGGACTTGCCATCTGGACCGAGTTCTAGCGCTTGCGTGAAG 475
Qy 407 AATACGTCATTCTCAAGGACCTGACCGAGAACCCGAGGCCACATGGCCGCTATGGAGA 466

RESULT 14
US-10-125-114-25069
; Sequence 25069, Application US/10425114
; Publication No. US200403488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5331)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 17144
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3689-245-C12_FLI
US-10-425-114-25069

Query Match Score 129.4; DB 13; Length 1289;
Best Local Similarity 51.8%; Pred. No. 9.5e-31; 0; Mismatches 356; Indels 15; Gaps 4;

Qy 107 TCACCACTGCCTGAAGACCCGTTACCTGACTTGGCTGTGACTCTGTAAC 166
Db 180 TCCACACAGCCTCCGTGCGTACGCCACCTGACTGGCCGTGACTACAGAAG 239
Qy 167 AGGGTGAGGTTGGTGGGGPATTCCTGACTTCCTGAAGAGAAACCCCTGGTGAAGCTG 226
Db 245 AAGCTGAGGTTGGTGAAGGGATTCAAGACGGACT -- GTCAAGGGCG 301
Qy 227 AGGACATCTTCTGACCAAGGTGTGACCAACTCCACCGTTATGAGGAACCTCCTCT 286
Db 302 AGGACCTGTTCATCACAAACCAAGCTGAGCTGAGCTAGACCTGCTG 355
Qy 287 GTCCATTGAGCTCCGAGGCTGTTGACTGAGTGTGATATGTTCTCGTCTG 346
Db 356 AAGGCTGCAAGAACCCGTTGAGGCTGAGCTGAGCTACCTCAAC 415
Qy 347 ACTGGCCATTTGCTGCCGAGAGATGCCGGGTGAGCCAGATGGCCCTGACGGA 406
Db 416 ATTCCTGAGGACTTGCCATCTGGACCGAGTTCTAGCGCTTGCGTGAAG 475
Qy 407 AATACGTCATTCTCAAGGACCTGACCGAGAACCCGAGGCCACATGGCCGCTATGGAGA 466


```

;FEATURE:
;    NAME/KEY: modified_base
;    LOCATION: (717)
;    OTHER INFORMATION: a, c, t, g, other or unknown
;FEATURE:
;    NAME/KEY: modified_base
;    LOCATION: (723)
;    OTHER INFORMATION: a, c, t, g, other or unknown
;FEATURE:
;    NAME/KEY: modified_base
;    LOCATION: (728)
;    OTHER INFORMATION: a, c, t, g, other or unknown
;FEATURE:
;    NAME/KEY: modified_base
;    LOCATION: (734)
;    OTHER INFORMATION: a, c, t, g, other or unknown
;US-10-004-115-19

Query Match          Score: 412: DB 15:
Best Local Similarity 100.0% ; Pred. No. 9.9-e-122
Matches 412; Conservative 0; Mismatches 0;

      567 CAACAGATCGAGATTCAACCCTTCCTGCCAACAGGAA
      1 CAACAGATCGAGATTCAACCCTTCCTGCCAACAGGAA
      627 CAAGACATATTGCCCGTGGCTACTCTCTCTGGGTCTC
      61  CAAGACATATTGCCCGTGGCTACTCTCTGGGTCTC
      QY          687 CGTGAGCCGGTCAAGAACAGACTTGACGATGAGAT
      121 CGTGAGCCGGTCAAGAACAGACTTGACGATGAGAT
      QY          747 CCTTGCTCAGGTTCTTATGCCCTGGGTCTGCCCGTCTG
      181 CCTTGCTCAGGTTCTTATGCCCTGGGTCTGCCCGTCTG
      Qy          807 CTCCAACCCAAAGGCCATTGAGTCAACTCAAGAGCT
      241 CTCCAACCCAAAGGCCATTGAGTCAACTCAAGAGCT
      Qy          867 TGAAGCCATTAAATGCCGTGGTCAAGGGTGTCACTTCG
      301 TGAAGCCATTAAATGCCGTGGTCAAGGGTGTCACTTCG
      Qy          927 TTTCGGATATGATGATGCTGCCGGAGAACGCCAAGAA
      361 TTTCGGATATGATGCTGCCGGAGAACGCCAAGAA

RESULT 8
US-10-369-493-27083
Sequence 2/083, Application US/10369493
; 1 Sequence 2/083, Application US/10369493
; 2 Publication No. US20030233675A1
; 3 GENERAL INFORMATION:
; 4 ; APPLICANT: Cao, Yongwei
; 5 ; APPLICANT: Hinkle, Gregory J.
; 6 ; APPLICANT: Slater, Steven C.
; 7 ; APPLICANT: Goldman, Barry S.
; 8 ; APPLICANT: Chen, Xianfeng
; 9 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEIN
; 10 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTY
; 11 ; FILE REFERENCE: 3:8-(52052) B
; 12 ; CURRENT APPLICATION NUMBER: US/10/369,493
; 13 ; CURRENT FILING DATE: 2003-02-28
; 14 ; PRIORITY APPLICATION NUMBER: US 60/3360,039
; 15 ; PRIORITY FILING DATE: 2002-02-21
; 16 ; NUMBER OF SEQ ID NOS: 47374
; 17 ; SEQ ID NO: 27083
; 18 ; LENGTH: 765
; 19 ; TYPE: DNA
; 20 ; ORGANISM: Neurospora crassa

```

Query Match		Best Local Similarity		Matches 549; Co	
Qy	46	CCTGG 	1	CCGGC 	
Db					
Qy	106	GTCAC 	61	GTCAA 	
Db					
Qy	166	GAGGG 	121	GAGGA 	
Db					
Qy	226	GAGCA 	181	GAGGA 	
Db					
Qy	286	TGGTC 			
Db					
Qy			241	TGGAG 	
Db					
Qy	346	CACTG 	301	CACTG 	
Db					
Qy	403	GGCAA 	361	GGCAA 	
Db					
Qy	463	GAGAA 			
Db					
Qy	523	GACCT 	481	GGCTT 	
Db					
Qy	583	CACCC 			
Db					
Qy	643	GTGGC 			
Db					
Qy	703	GAGGA 	601	ATTCG 	
Db					
Qy	763	ATTGC 			
Db					

QY 605 AGCTGGAGTACTGCTTCAGAACATTATGCCGTGCTACTCTCCCTGG 661
 Db 361 AGCTGTGCACTGCTTCAGAACATTATGCCGTGCTACTCTCCCTGG 417

RESULT 7

GENERAL INFORMATION:
 / Publication No. US20030134402A1
 / APPLICANT: ASAKO, HIROYUKI
 / APPLICANT: MATSUMURA, KENJI
 / APPLICANT: SHIMIZU, MASATOSHI
 / APPLICANT: ITO, NORIBU
 / APPLICANT: WAKITA, RYUEI

TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
 FILE REFERENCE: 7-2-7249
 CURRENT APPLICATION NUMBER: US/10/004,115A
 PRIOR APPLICATION NUMBER: JP 2000-372704
 PRIOR FILING DATE: 2000-12-07
 PRIOR APPLICATION NUMBER: JP 2001-006144
 PRIOR FILING DATE: 2001-01-15
 PRIOR APPLICATION NUMBER: JP 2001-026594
 PRIOR FILING DATE: 2001-02-02
 PRIOR APPLICATION NUMBER: JP 2001-175175
 PRIOR FILING DATE: 2001-06-11
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 19
 LENGTH: 743
 TYPE: DNA
 ORGANISM: Escherichia coli

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (434)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (440)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (443)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (448)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (458)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (475)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (506)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (510)

FEATURE: OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (517)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (522)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (533)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (583) . (584)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (607)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (616)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (634)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (639)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (648)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (651) . (652)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (659)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (662)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (670) . (671)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (690)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (693)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (695)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (702)

OTHER INFORMATION: a, c, t, g, other or unknown
 NAME/KEY: modified_base
 LOCATION: (707)

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (463)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (485)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (489)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (491)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (520)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (524)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (593)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (602)
OTHER INFORMATION: a, c, t, g, other or unknown

FEATURE: modified_base
NAME/KEY: modified_base
LOCATION: (682)
OTHER INFORMATION: a, c, t, g, other or unknown

US-10-004-115A-15

Query Match Score 588.6; DB 15; Length 697;
Best Local Similarity 94.4%; Pred. No. 1.5e-178;
Matches 611; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Qy 4 TCTAAAGGAAAGACTTCAATGAGCAACGGCTAACATTCTGGCCGCTGGTTGGT 63
Db 52 TCCRAAGGAANACTTCACACTGAGCAGGGCTCAAAAATTCTGGGTGGTTGGT 111
Qy 64 ACCCTTGCTACTGAGGTTCAAGGGGAGACCTTAATCTGTCAGCCTGAG 123
Db 112 ACCTNGCTGTGAAGTTCAAGGGGAACCTATNCCTGTCACACTGCCATAA 171
Qy 124 ACCGGTTACCGTCACTTGGACTGCTGCTGAGCTGAGGGTGAACGGGTTGGTGGAG 183
Db 172 ACCGGTTACCGTCACTTGGACTGCTGCTGAGCTAAGGGTGAAGTTGGTGGAG 231
Qy 184 GGATCCGCTGACTCTGAGGAGAACCCCTCGTGAAGCTGAGGAATCTCTGTC 243
Db 232 GGTTTCGGTGAACCTTCCTGAGGGAAACCCCTCGTGAAGCTGAGGAATCTCTGTC 291
Qy 244 ACCAAGGTGTTGGACTTGTGACTACGGTCCATGAGGACTCC 303
Db 292 ACCAAGGTGGAACCCCTCCAGGTTCAACTGGCTTCAATGAGGATCTCC 351
Qy 304 CTGAAGCGTCTGGACTTGTGACTACGGTCCATGTTGAGGACATCTGTC 363
Db 352 CTGAAGCGTCTGGACTTGTGACTACGGTCCATGTTGAGGACATCTGTC 411
Qy 364 GAGAGAAATGCCAGGTGAGCCCAAGATTGGCCAAATACGTCACTCTCAG 423
Db 412 GAAAAAAATGGCAGGGTGCCAAATTTGGCTTGAGCAATACNTCTCAG 471
Qy 424 GACCTGACCGAGAACCCGGCCACATGGCCGCTATGGAGAAAGATTATGAGGATGC 483
Db 472 GACCTGACCGAGAACCCCNACCCACTGGCGCTATGGAAAATTTNGANGATCC 530

RESULT 6
US-10-004-115A-22
i Sequence 22, Application US/10004115A
i Publication No. US20030134402A1
i GENERAL INFORMATION:
i APPLICANT: ASAKO, HIRYUKI
i APPLICANT: MATSUMURA, KENJI
i APPLICANT: SHIMIZU, MASATO
i APPLICANT: ITO, NORIYA
i APPLICANT: WAKITA, RYUHEI
i TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
i 4-HALO-3-HYDROXYBUTANOATE
i FILE REFERENCE: 7372-7249
i CURRENT APPLICATION NUMBER: US/10/004,115A
i CURRENT FILING DATE: 2002-10-23
i PRIOR APPLICATION NUMBER: JP 2000-372704
i PRIOR FILING DATE: 2000-12-07
i PRIOR APPLICATION NUMBER: JP 2001-006144
i PRIOR FILING DATE: 2001-01-15
i PRIOR APPLICATION NUMBER: JP 2001-026594
i PRIOR FILING DATE: 2001-02-02
i PRIOR APPLICATION NUMBER: JP 2001-175175
i PRIOR FILING DATE: 2001-06-11
i NUMBER OF SEQ ID NOS: 37
i SOFTWARE: PatentIn Ver. 2.1
i SEQ ID NO: 22
i LENGTH: 417
i TYPE: DNA
i ORGANISM: Escherichia coli
i US-10-004-115A-22

Query Match Score 42.6%; DB 15; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.9e-123;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 245 CCAAGGTGTGSAACACTCCACCTTATAGGAGTGTCTCTGTCCATTGACACTCC 304
Db 1 CCAAGGTGTGSAACACTCCACCTTATAGGAGTGTCTCTGTCCATTGACACTCC 60
Qy 305 TGAAGGCCTCTGGACTTGTACAGTGTGATGTCTCGTTCACTGGCCATTGTGCCG 364
Db 61 TGAAGGCCTCTGGACTTGTACAGTGTGATGTCTCGTTCACTGGCCATTGTGCCG 120
Qy 365 AGAAGATGGCAGGGTGAACCTGAGCCAAATAGCTGATTCTCAAGG 424
Db 121 AGAAGATGGCAGGGTGAACCTGAGCCAAATAGCTGATTCTCAAGG 180
Qy 425 ACCTGACCGAGAACCCGGCCACATGGGGCTATGGAGAAATTTATGAGGATCGCA 484
Db 181 ACCTGACCGAGAACCCGGCCACATGGGGCTATGGAGAAATTTATGAGGATCGCA 240
Qy 485 AGGCCAGGTCTTGTGCTCATGGTCTCCACTGGCCTGAGGAGATGTCGCAAGT 544
Db 241 AGGCCAGGTCTTGTGCTCATGGTCTCCACTGGCCTGAGGAGATGTCGCAAGT 300
Qy 545 TGGCCAAAGGTATGGCTACGCCAACAGTCGAGATCACCCCTCTGCCAACAGG 604
Db 301 TGGCCAAAGGTATGGCTACGCCAACAGTCGAGATCACCCCTCTGCCAACAGG 360

PRIOR FILING DATE: 2001-06-11
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 28
 LENGTH: 996
 TYPE: DNA
 ORGANISM: Penicillium citrinum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1) .. (975)
 US-10-004-115A-28

Query Match 100.0%; Score 978; DB 15; Length 996;
 Best Local Similarity 100.0%; Pred. No. 1e-303; Gaps 0;
 Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCTAACGGAAAGACTTACATTGAGCAACCGGTCAAAGATTCCGGTGGCTTT 60
 Db 1 ATGTCCTAACGGAAAGACTTACATTGAGCAACCGGTCAAAGATTCCGGTGGCTTT 60
 QY 61 GGTAACCTTCCCTACTGAAAGTTCCAAGGGGAGACCTATACTGTGTCAACCCTGCTG 120
 Db 61 GGTAACCTTCCCTACTGAAAGTTCCAAGGGGAGACCTATACTGTGTCAACCCTGCTG 120
 QY 121 AGACCGGTAAACGGTCACTTGACTGTTGCTGGTACTACTGTGAGGGTAGGGTTGT 180
 Db 121 AGACCGGTAAACGGTCACTTGACTGTTGCTGGTACTACTGTGAGGGTAGGGTTGT 180
 Qy 181 GAGGTATCGTGAATTCTCGTAGGAAACCCCTGGTAGGACATTTCTCTC 240
 Db 181 GAGGTATCGTGAATTCTCGTAGGAAACCCCTGGTAGGACATTTCTCTC 240
 Qy 241 TGCACCAAGGTGGAAACACCTCCACCGTTATGGAGCTCTGGCCATTGACC 300
 Db 241 TGCACCAAGGTGGAAACACCTCCACCGTTATGGAGCTCTGGCCATTGACC 300
 Qy 301 TCCCTGAAGCTCTCGACATGCTACGTTGATGTTGCTCGTCACTGGCCATTGCT 360
 Db 301 TCCCTGAAGCTCTCGACATGCTACGTTGATGTTGCTCGTCACTGGCCATTGCT 360
 Qy 361 GCGGAAAGATGGCCAGGTTGAGCCAAAGATTGCCGCAATTACGTCATCTC 420
 Db 361 GCGGAAAGATGGCCAGGTTGAGCCAAAGATTGCCGCAATTACGTCATCTC 420
 Qy 421 AAGGACTCTACCGAAACCCGGAGCCAAATGGCGGCTATGGGAAGTTATGAGAT 480
 Db 421 AAGGACTCTACCGAAACCCGGAGCCAAATGGCGGCTATGGGAAGTTATGAGAT 480
 Qy 481 CGCAAGGCTGGTCCATTGGTCTCOAATGGCCATTGCGGACTTGGAGAGATGTC 540
 Db 481 CGCAAGGCTGGTCCATTGGTCTCOAATGGCCATTGCGGACTTGGAGAGATGTC 540
 Qy 541 AAGTTCGCAAGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGG 600
 Db 541 AAGTTCGCAAGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGG 600
 Qy 601 GAGGAGCTGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGG 660
 Db 601 GAGGAGCTGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGGTCACTGG 660
 Qy 721 GAGATGGCTGGAGAAGGGGAAACCCCTGCAAGATTTCAAAACATTATGCGTGGCTT 780
 Db 721 GAGATGGCTGGAGAAGGGGAAACCCCTGCAAGATTTCAAAACATTATGCGTGGCTT 780
 Qy 781 CGTGGCTGAAAGAACCGGTTCACCCACGGTCAACCTGAGGAAACAGACTCTGAAAC 720
 Db 661 GGCTCGAGAACCGATTTCCACCCACGGTCAACCTGAGGAAACAGACTCTGAAAC 720
 Qy 781 CGTGGCTGAAAGAACCGGTTCACCCACGGTCAACCTGAGGAAACAGACTCTGAAAC 840
 Db 781 CGTGGCTGAAAGAACCGGTTCACCCACGGTCAACCTGAGGAAACAGACTCTGAAAC 840
 Qy 841 AGCATTCGGTCTCCSATGGCAACTTGTGCAAGGGTGTGCAAACTCGTTGCAAGGGTGTGCA 900

RESULT 5
 US-10-004-115A-15
 ; Sequence 15, Application US/10004115A
 ; Publication No. US20030134402A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ASAKO, HIROYUKI
 ; APPLICANT: MATSUMURA, KENJI
 ; APPLICANT: SHIMIZU, MASATOHI
 ; APPLICANT: ITO, NOBUTHI
 ; APPLICANT: WAKITA, RYUHEI
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE
 ; TITLE OF INVENTION: 4-HALO-3-HYDROXYBUTANOATE
 ; FILE REFERENCE: P372-7219
 ; CURRENT APPLICATION NUMBER: US/10/004,115A
 ; CURRENT FILING DATE: 2002-10-23
 ; PRIOR APPLICATION NUMBER: JP 2000-372704
 ; PRIOR FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: JP 2001-006144
 ; PRIOR FILING DATE: 2001-01-15
 ; PRIOR APPLICATION NUMBER: JP 2001-025594
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: JP 2001-175175
 ; PRIOR FILING DATE: 2001-06-11
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 697
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (14)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (53)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (116)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (148)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (235)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (185)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (346)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (460)
 ; OTHER INFORMATION: a, c, t, g, other or unknown

601	GAGGGAGTGGTGCAGTACTGCTTCCTCAAGAACATTATGCCGTCCCTACTCTCTG	660
661	GGCTCGAGAACCGTTCCACACCGGTAGGGTCAAGAGAACAGACTCTGAAAC	720
661	GGCTCGAGAACCGGTTCCACACCGTCAAGAGAACAGACTCTGAAAC	720
721	GAGATGCCGAGAACGGGGCAACACCCCTGGCTCAGCTCTTATTGGCTGGGTCTGGC	780
721	GAGATGCCGAGAACGGGGAAACACCCCTGGCTCAGCTCTTATTGGCTGGGTCTGGC	780
781	CCTGGCTAAGCTCGTCTCCCAAGGAGCTCAAACCCAAAGGGCATGGAGTCCAACATCAAG	840
781	CCTGGCTAAGCTCGTCTCCCAAGGAGCTCAAACCCAAAGGCCATTAGTCCAACATCAAG	840
841	AGCATTAAGGTCTGGATGCCGACTTGTAGGCCATTAATGCCGTTCAGGGTCTTCAC	900
841	AGCATTAAGGTCTGGATGCCGACTTGTAGGCCATTAATGCCGTTCAGGGTCTTCAC	900
901	TTCCCTTTCGTCATACTGAGGATACTTTCGATATGATGTCTGGCCGAGGAGACCGCC	960
901	TTCCCTTTCGTCATACTGAGGATACTTTCGATATGATGTCTGGCCGAGGAGACCGCC	960
961	AAGAACCTGTCGCTGA	978
961	AAGAACCTGTCGCTGA	978

RESULT 3
US-10-608-533-2
; Sequence 2, Application US/10608533
; Publication No. US20040091902A1
; GENERAL INFORMATION:

Query Match 100.0%; Score 978; DB 17; Length 978;
 Best Local Similarity 100.0%; Pred. No. 1e-303; Gaps 0;
 Matches 978; Conservative 0; Mismatches 0; Indels 0;

US-10-608-533-2

QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db
1	1	61	61	121	121	181	181	241	241	QY	
ATGTCATTACGGAAAGACTTTCACATTGAGCAACGGGTCAAGATTCCTGGGGTGCCTT	ATGTCATTACGGAAAGACTTTCACATTGAGCAACGGGTCAAGATTCCTGGGGTGCCTT	GGTACCTTCGCTTAGTGAAAGTTCCAAAGGGCGAGACCTTAATCTGCTGTCACCAACTGCCCTG	GGTACCTTCGCTTAGTGAAAGTTCCAAAGGGCGAGACCTTAATCTGCTGTCACCAACTGCCCTG	AAGCCGGTACCGTACCTGGACTCCCTGGTGAATACCTGAACAGGGTGAGGTGTT	AAGCCGGTACCGTACCTGGACTCCCTGGTGAATACCTGAACAGGGTGAGGTGTT	GAGGTATCCGTGACTTCCCTGAAGGAAACCCCTCCGTGAAGCGTGAGGACATCTCGTC	GAGGTATCCGTGACTTCCCTGAAGGAAACCCCTCCGTGAAGCGTGAGGACATCTCGTC	TGCAAGGGTACCCGTTACCGTACCTGGACTCCCTGGTGAATACCTGAACAGGGT	TGCAAGGGTACCCGTTACCGTACCTGGACTCCCTGGTGAATACCTGAACAGGGT	;	
60	60	120	120	180	180	240	240	300		;	
978	978	121	121	181	181	241	241	300		;	
(1) .. (975)										LOCATION:	NAME/KEY:
										FEATURE:	ORGANISM:
										PENICILLIUM CILIINUM	

Db	241	TGGACCAAGGTGTTGAAACCACTCCACCGTTATGAGCAGCTCCATTGGTCATGACGAC	300
Qy	301	TCCCTGAAGCTCTTGCAGTGAACGTGATATGTTCTCGTCACTGGCCCATATGCT	360
Db	301	TCCCTGAAGCTCTTGCAGTGAACGTGATATGTTCTCGTCACTGGCCCATATGCT	360
Qy	361	GCGGAGAAGAATGGCGGGTAGGCCAAGATGGGCTTGACGGAAATAGTCATCTC	420
Db	361	GCGGAGAAGAATGGCGGGTAGGCCAAGATGGGCTTGACGGAAATAGTCATCTC	420
Qy	421	AAGGACCTGACCGAACCCCCGAGGCCACATGGCCCTATGGAGAGATTATGGAGAT	480
Db	421	AAGGACCTGACCGAACCCCCGAGGCCACATGGCCCTATGGAGAGATTATGGAGAT	480
Qy	481	CGCAAAGGCCAGSTCCATTGGTCTCAACTGGACCATTCGAGACTGTGCC	540
Db	481	CGCAAAGGCCAGSTCCATTGGTCTCAACTGGACCATTCGAGACTGTGCC	540
Qy	541	AAGTTGCCAAGGTCTGGCTAACGGCAACCATGGAGATTCAACCCCTCCGCCAAC	600
Db	541	AAGTTGCCAAGGTCTGGCTAACGGCAACCATGGAGATTCAACCCCTCCGCCAAC	600
Qy	601	GAGGAGTGGTCACTGAGFACTGCTTCAGAACATTAGCCCGTGGCFACTCTCTG	660
Db	601	GAGGAGTGGTCACTGAGFACTGCTTCAGAACATTAGCCCGTGGCFACTCTCTG	660
Qy	661	GGCTCGCGAACCGGTTCCACCAACCGGTCAAGGGTCAAGAACATTAGCTGAA	720
Db	661	GGCTCGCGAACCGGTTCCACCAACCGGTCAAGGGTCAAGAACATTAGCTGAA	720
Qy	721	GAGATCGCGAACGGGGCAACCCCTGGTCAACGGCAACCCAGCTTCAACTTCAAG	780
Db	721	GAGATCGCGAACGGGGCAACCCCTGGTCAACGGCAACCCAGCTTCAACTTCAAG	780
Qy	781	CGTGGCTAGCTGTCTCCCAAGAGCTTCAACCCAGCTTCAACCCAGCTTCAACTTCAAG	840
Db	781	CGTGGCTAGCTGTCTCCCAAGAGCTTCAACCCAGCTTCAACCCAGCTTCAACTTCAAG	840
Qy	841	AGGATTAAGCTTCCGATGCCGACTTGTGAAGCCATAATGCCGTGCAAGGGTGTCA	900
Db	841	AGGATTAAGCTTCCGATGCCGACTTGTGAAGCCATAATGCCGTGCAAGGGTGTCA	900
Qy	901	TCGGTTCTGTCAACTGAAGGATATTTCGGATATGATCTGCCCGAGGGAGACGCC	960
Db	901	TCGGTTCTGTCAACTGAAGGATATTTCGGATATGATCTGCCCGAGGGAGACGCC	960
Qy	961	AAGAACCTGTCGGCTCA	978
Db	961	AAGAACCTGTCGGCTCA	978

RESULT 4
US-10-004
; Sequence
; Publica
; GENERAL
; APPLIC
; APPLIC
; APPLIC
; APPLIC
; APPLIC
; TITLE
; TITLE
; FILE R
; CURRENT
; CURRENT
; PRIORITY
; PRIORITY
; PRIORITY
; PRIORITY
; PRIORITY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2004, 11:34:38 ; Search time 495 Seconds
(without alignments)
8987.526 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978

Sequence: 1 atgtccaacggaaaggacttt.....ccaaagaacctgtctgcgtga 978

Scoring table: IDENTITY_NUC
Gapop 1.0 , Gapext 1.0

Searched: 2980401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters: 5920802

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
 1: /cgnd2_6/podatata/1/pubnpna/US07_PUBCOMB.seq:
 2: /cgnd2_6/podatata/1/pubnpna/US07_PCT_NEW_PUB.seq:
 3: /cgnd2_6/podatata/1/pubnpna/US07_NEW_PUB.seq:
 4: /cgnd2_6/podatata/1/pubnpna/US07_PUBCOMB.seq:
 5: /cgnd2_6/podatata/1/pubnpna/US07_NEW_PUB.seq:
 6: /cgnd2_6/podatata/1/pubnpna/PECTUS_PUBCOMB.seq:
 7: /cgnd2_6/podatata/1/pubnpna/US08_PUBCOMB.seq:
 8: /cgnd2_6/podatata/1/pubnpna/US08_PUBCOMB.seq:
 9: /cgnd2_6/podatata/1/pubnpna/US09A_PUBCOMB.seq:
 10: /cen2_6/podatata/1/pubnpna/US09B_PUBCOMB.seq:
 11: /cen2_6/podatata/1/pubnpna/US09C_PUBCOMB.seq:
 12: /cen2_6/podatata/1/pubnpna/US09_NEW_PUB.seq:
 13: /cen2_6/podatata/1/pubnpna/US09_NEW_PUB.seq:
 14: /cen2_6/podatata/1/pubnpna/US10A_PUBCOMB.seq:
 15: /cen2_6/podatata/1/pubnpna/US10B_PUBCOMB.seq:
 16: /cen2_6/podatata/1/pubnpna/US10C_PUBCOMB.seq:
 17: /cen2_6/podatata/1/pubnpna/US10_NEW_PUB.seq:
 18: /cen2_6/podatata/1/pubnpna/US60_NEW_PUB.seq:
 19: /cgnd2_6/podatata/1/pubnpna/US08_PUBCOMB.seq:
 Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	978	100.0	978	15	US-10-004-115A-2
2	978	100.0	978	15	US-10-327-108-4
3	978	100.0	978	17	US-10-608-533-2
4	978	100.0	996	15	US-10-004-115A-28
5	588.6	60.2	697	15	US-10-004-115A-15
6	7	41.7	417	15	US-10-004-115A-22
7	7	41.2	42.1	15	US-10-004-115A-19
8	408.2	41.7	765	16	US-10-364-493-27083
9	265	27.1	331	15	US-10-004-115A-18
10	199.6	20.4	972	13	US-09-734-237B-31
11	144.2	14.7	1121	14	US-10-04-416-5
12	142.6	14.6	1077	14	US-10-040-416-3
13	130.8	13.4	1289	13	Sequence 3, Appli
14	129.4	13.2	1289	13	Sequence 17144, A

Sequence 25069, A

Sequence 30950, A

Sequence 20425, A

Sequence 1010, Ap

Sequence 332, App

Sequence 641, App

Sequence 641, App

Sequence 1, Appli

Sequence 1, Appli

Sequence 43, Appli

Sequence 29, Appli

Sequence 1492, Ap

Sequence 12, Appli

Sequence 47, Appli

Sequence 11, Appli

Sequence 16, Appli

Sequence 8, Appli

Sequence 12327, A

Sequence 5041, Ap

Sequence 15192, A

Sequence 12, Appli

Sequence 15, Appli

Sequence 85, Appli

Sequence 10, Appli

Sequence 91, Appli

Sequence 150, App

Sequence 147, App

Sequence 33, Appli

Sequence 434, App

Sequence 434, App

Sequence 434, App

Sequence 434, App

ALIGNMENTS

RESULT 1
US-10-004-115A-2

Sequence 2, Application US/10004115A
Publication No. US20030134402A1
GENERAL INFORMATION:
APPLICANT: ASAKO, HIROYUKI
APPLICANT: MATSUMURA, KENJI
APPLICANT: SHIMIZU, MASATO SHI
APPLICANT: ITO, NOBUHI
APPLICANT: WAKITA, RYUHEI
TITLE OF INVENTION: PROCESS FOR PRODUCING OPTICALLY ACTIVE 4-HALO-3-HYDROXYBUTANOATE
FILE REFERENCE: 7372-72249
CURRENT APPLICATION NUMBER: US/10/004,115A
CURRENT FILING DATE: 2002-10-23
PRIOR APPLICATION NUMBER: JP 2000-372704
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: JP 2001-006144
PRIOR FILING DATE: 2001-01-15
PRIOR APPLICATION NUMBER: JP 2001-026594
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: JP 2001-175175
PRIOR FILING DATE: 2001-06-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 978
TYPE: DNA
ORGANISM: Penicillium citrinum
FEATURE: CDS
NAME/KEY: CDS
LOCATION: (1)..(975)

US-10-004-115A-2

Query Match Score 978; DB 15; Length 978;
Best Local Similarity 100.0%; Prd. No. 1e-303;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	323	ACTACGTTGATATGTCCTCGTCACTGGCCATTGCTCGAGAAGATGGCCAGGGTG	382
Db	380	AGTACCTCGATCTGTAACCTPATCCACTGGCCATGGCTAACGGAG-----G	427
Qy	383	AGCCCAAGATGGCCCTGACGCCAATACTGCATCTCAAGGACCTGAGAACCCG	442
Db	428	GATGCCATCTGTCCTCCACGCCACAGGTGCAAGACGCTGTAATCGCCGGTATTACG	487
Qy	443	AGCCCAATGGCCCTATGAGGATTATGAGCATGGCAAGGCGAGTCATTGCTG	502
Db	488	TCGAACCTGTAAGGCCATGAGAAGTTCGGTCAAGTCCATTGCTG	547
Qy	503	TCTCCAACCTGGACCATGGCACTTGAGAAGATGCCAAGGTATGCCCTC	562
Db	548	TTTCCAACCTCAACGAAGCCAGTCGGCGCTCTGAGAAGATGCCAAGGTATCCACAG	607
Qy	563	ACGCCAACCGAGATGGATTACCCCTTCTGCCAACGGAACTGTCGAGTACTCT	622
Db	608	TAACCCRATGGATGGTCGATCCATATGACCCAGAGCTGATTGACTTCTGCA	667
Qy	623	TCTCCAAAGAACATTATGCCGCGCTACTCTCTGCGCTCCAGAACAGGTTCGA	682
Db	668	AGTCAAAGGACATTAACATACTGCTAACGTCATCTGATCTCCACCGCCATGGCA	727
Qy	683	CCACCGGTGAGCGGTCAAGGAGAACAG	711
Db	728	AGGCTGGTGTACGTCATAGCCAGGG	756

Search completed: May 29, 2004, 12:59:21
 Job time : 2529 secs

JOURNAL	Unpublished (2002)	Qy	596 CCACAGGAGCTGGTCACTGCCTACTGCCATTATGCCGTCCTACTCTC 655
COMMENT	Contact: Stein Nils Molecular Markers Group, Department Genbank Institute of Plant Genetics and Crop Plant Research (IPK) Corrensstr. 3, 06466, Gatersleben, Germany	Db	552 TCCAGGAGGTCACTGCCTACTGCCATTATGCCGTCCTACTCTC 611
	Tel: 039482-5522 Fax: 039482-5595	Qy	656 CTCTGGGC 663
	Email: stein@ipk-gatersleben.de	Db	612 CCTTCGGC 619
FEATURES		RESULT 15	
source		AA697426	758 bp mRNA linear EST 19-APR-2001
Location/Qualifiers		LOCUS	HL02448 5prime HL Drosophila melanogaster head BlueScript mRNA sequence.
1..638	/organism="Hordeum vulgare subsp. vulgare" /mol_type="mRNA" /cultivar="barke" /sub_species="vulgare" /db_xref="CAPI:26.815" /db_xref=taxon:112509 /clone=HT13n05" /issue_type="endosperm early" /dev_stage="0-16 hours after imbibition" /lab_host="X110-Gold" /clone_lib="HT" /note="Vector: pBlueScript SK+; Site_1: EcoRI (5'-end of cDNA); Site_2: XbaI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, Sall, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."	DEFINITION	Drosophila melanogaster cDNA clone HL2448 5prime, mRNA sequence.
		ACCESSION	AA697426
		VERSION	AA697426.1 GI:2700355
		KEYWORDS	EST.
		SOURCE	Drosophila melanogaster (fruit fly)
		ORGANISM	Drosophila melanogaster
		REFERENCE	Eukaryota; Marazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Dictyoptera; Blattodea; Muscomorpha; Ephydioidea; Drosophilidae; Drosophila.
		AUTHORS	Harvey,D., Brookstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.
		TITLE	BDGP HHMI Drosophila EST Project
		JOURNAL	Unpublished (2001)
		COMMENT	Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab, Berkeley, CA 94720, USA One Cyclotron Rd, Berkeley, CA 94720, USA Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Fax: 510 486 6798
			Place: 24 row: D column: 12 High quality sequence stop: 630.
		FEATURES	Location/Qualifiers
		source	1..758 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="HU2448" /sex="male and female" /dev_stage="adult" /lab_host="SOL" /clone.lib="HL Drosophila melanogaster head BlueScript" /note="Organ: head-brain & sensory organ; Vector: BlueScript SK; Site_1: EcoRI Site_2: XbaI; Constructed using Stratagene ZAP-cDNA Synthesis Kit. Oligo dN-primer and directionally cloned at EcoRI and XbaI in BlueScript SK (+/-)"
ORIGIN		ORIGIN	
Query Match	Score 148.8; DB 13; Length 638;	Query Match	Score 148.6; DB 9; Length 758;
Best Local Similarity	55.1%; Pred. No. 3.6e-22;	Best Local Similarity	54.8%; Pred. No. 4.4e-22;
Matches	335; Conservative 0; Mismatches 267; Indels 6; Gaps 2;	Matches	345; Conservative 0; Mismatches 269; Indels 8 15; Gaps 2;
Qy	59 TTGGTACCTCGCTAATGAAAGTCCAAGGGAGACCTATACTGCTGACCATGCC 118	Qy	83 CCAGGGAGACCTATACTGCTGACCATGCCATTACGGTACCTGG 142
Db	15 TTGGTTCGGTACATGGAAAGCTGAGCTGGCTGCTGGCTGAGGCC 74	Db	143 CCAGGGTCAAGTCAGCTGGCTGAAAGTGGCTGAGCTGGCTGAGGCCATTG 202
Qy	119 TGAAGAACGGTACCGTCACTGGACTGTGCTGGTACACTGTAACGGGGTGAGGTG 178	Db	143 ACTGGCTGCTGTTACTCTGAACTGGGTGAGGTGAGGTGACTCTGTA 202
Db	75 TCAAGCTGGTACCCACCTTGTATCTGGCAAGGTACGGAACAGATG 134	Qy	203 ACCTGGCTTATGGAGCTCCCTGGTCACTGGACATGGTGGTGGTGG 259
Qy	179 GTGAGGTTATCGTGCAC- -TTCCCTGAAGGAGAACTCCCTGGTAGGTACATC 235	Db	203 AGAGAACCCCTGGTGAAGCTGGTCACTGGTGGTGGTGG 262
Db	135 CCGGGCTCTTAAAGGCTTGGGGGGCTTACCTCAAGCTGAGACGTG 194	Qy	203 AGAGAACCCCTGGTGAAGCTGGTCACTGGTGGTGGTGG 262
Qy	236 TCGTCTGACCAAGGCTGGAAACCRACCTCCACCGTTATGAGGACTCCCTGGTCCATG 295	Db	260 TCAAGGGGGCTTGTCAAGCTGGTGGTCACTGGACATGGTGGTGGTGG 319
Db	195 TCATCACATCCAAAGCTGGACTCCAGAACGCCCTCTG 254	Qy	263 TCCACCGTTATGGAGCTCCCTGGTCACTGGACATGGTGGTGGTGGTGG 322
Qy	296 ACGACTCCCTGAAGCGCTTGGACTCTGATGTTCTGTTCACTGCCA 355	Db	320 TCCATGCCGCGGCTCACTCCGGTCAACGGACCTCCGGTCAACGGACCTCCGG 379
Db	255 ACGACTGCTGCTGGCTGAGCTTGGCTGACCTAGACCTTACCTGTCATCTCCCG 314	Qy	535 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491
Qy	356 TTGGCTGCGAGAAAGTGGCAAGGGTGAACCTGGCTACGGAAATACTGCA 415	Db	535 AGAGGAACTGGCTGAGGACCTGGCTACGGACCTGGCTACGGACCTGGCTACGG 551
Db	315 TCGCTTGTGACCGAGCGGGGAGGAGC 374	Qy	536 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491
Qy	416 TTCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 475	Db	536 AGAGGAACTGGCTGAGGACCTGGCTACGGACCTGGCTACGGACCTGGCTACGG 551
Db	375 TCAAGTGTCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 434	Qy	537 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491
Qy	476 AGGATCGAAGGGCACTCCATTGCTGACCTGGCTACCTGGCTACGGACCTGGCTACGG 535	Db	537 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491
Db	435 AGGA--GAAGGCGGGCTGCTGGCTACGGACCTGGCTACGGACCTGGCTACGG 535	Qy	538 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491
Qy	536 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491	Db	538 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491
Db	492 TCTCGAGGACCCGGCTCACTCCGGTCACTGGCTACGGACCTGGCTACGG 551	Qy	539 TCTCAAGGACCTGACCGAACCCCACATGGCCATGAGCCATGGTGG 491

source 1. .189
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="hhol01"
 /tissue_type="vegetative mycelia, asexual structures"
 /clone_libr="Aspergillus nidulans 24hr asexual"
 /note="vector: pBlueScript SK-; Site 1: EcoRI; Site 2:
 XbaI; 5' end of cDNA cloned into EcoRI site of pBlueScript
 3', end of cDNA cloned into XbaI site of pBlueScript"

ORIGIN

Query Match	Match	Score	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
	Matches	331	Conservative	0;	3.6e-22;		0;	245	2;
Qy	732	GAAGGGCGAAACCCCTGCTCAGGTTCTATGCCCGGGTGCCTGT	15.2%; Length 189; Score 152.2; DB 9;	83	CCAAGGGCGAGACCTATACGTGTCAACACTGCTGAAGACCGGTACGGTACTGG	142			
Db	1	GAAGGGCGAAACCCCTCCCTAGSTCTCATGGCTGGCTGCTACGT	87.8%; Pred. No. 3.3e-23; Matches 166; Conservative 0; Mismatches 23;	143	CCAAGGGCTGGTCAGGGCTGTCAAGTGTGATGTGCGGATACAGGGACATGG	202			
Qy	792	CGTTCTCCCAAAGCTCCAAACCCCAAGGCCATTGAGTCACACTCAAGGATGGACT	14.3%; Length 180; Score 152.2; DB 9;	143	ACTGTGCTGGTACTACTGTGACAGGGTGAGGTGGTGGGGTACGTTCCTGTA	202			
Db	61	CGTCCTGCCAACAGCTAACCCGCTGCAACCGGGCATGATGCAACTTCAGAACATCGAGCT	85.1%; Pred. No. 3.3e-23; Matches 166; Conservative 0; Mismatches 23;	203	ACTGTGCTTAATGATACAAACAGGATAAGGGATGGCTGGATGGTCT--TGAGGCCAAGA	259			
Qy	852	CTCCCATGCGGACTTGTGACCCATAATGCCGTTGCAAGGGTGTGCAACTTCCGTTCTGCT	14.3%; Length 180; Score 152.2; DB 9;	203	AGGAGAACCCCTGGTGAAGCGTGGAGGACATCTCTGACCAAGGGTGGACACC	262			
Db	121	GTCCCACGAGCACTAGAGCCGCTAACGGCCGTAAACGGTGTGCAAGGGTCACTCCGCTGT	85.1%; Pred. No. 3.3e-23; Matches 166; Conservative 0; Mismatches 23;	260	TCAAGGGGGGGTGTCAAGTGGATGGATGTGTCATCACAGAACATCTGGGACACTT	319			
Qy	912	CAACATGAA 920	14.3%; Length 189; Score 152.2; DB 9;	263	TCCACCGTTATGAGGAGCTCTCTGGTCCATTGAGGACTCCCTGAAGGGCTCTGGACTG	322			
Db	181	CAACATGAA 189	14.3%; Length 189; Score 152.2; DB 9;	320	TCCATGCCGATCTGTCAGTGGCATTTGGAGAACACATGAGCTCCTGAGCTA	379			
Qy				323	ACTACGTGATATGTTCTCTCTGCTCTGGCCATGTCAGGAAATGGCCAGGGTG	382			
Db				380	AGTACCTCGATCTGTACCTPATGCCATGGCCATGGCTACAAGGAG-----G	427			
Qy				383	AGGCCAGATGGCTCTCTGGCAATACTCTCAGGACTCTGTCAGGAACTGCAAGG	442			
Db				428	GATGCCATCTGTTCCACCCGACAGGATGGCAAGACGCTGTACTCGGGTGTGATTAGC	487			
Qy				443	AGCCCACATGGCCATTGAGGAGATTATGAGGATCCAAAGCCAGTCCATTGGTG	502			
Db				488	TGACAGCTGAAAGGCACTGAGANTGTGTGGAGAGGAGGTGTCAAGTCCATTGGTG	547			
Qy				503	TCTCCAACTGACCATTGCCGACCTTCAAGATGTCAGGACTCTGACCGAGCTC	562			
Db				548	TTTCCCACTTAAACAGAGGAGATGAGGGCTCTGAGGCTCCACTATTCCACAG	607			
Qy				563	ACGCCAACAGATCGAGATTACCCCTCTGCCCCAACAGGAGCTGGCAGTACTGCT	622			
Db				608	TAACCCATCACAGATGTCGCACTCCATATCTGACCCAGAAAGCTGATCTGCA	667			
Qy				623	TCTCCAGAAATTGCCCCGGCTACTCTCTGGCTCGAGAAC	673			
Db				668	AGTCACAGGACATTACATCACTGCTAACGGCTTGGGATCTCCCAACC	718			
COMMENT					RESULT 14				
KEYWORDS					CA015279				
ORGANISM	Drosophila melanogaster				LOCUS	HT13N05r			
DEFINITION	Drosophila melanogaster cDNA clone LD063335 similar to CG0084:				DEFINITION	HT13N05r HT Hordeum vulgare subsp. vulgare cDNA clone HT13N05			
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				ACCESSION	CA015279			
VERSION	AA263299				VERSION	5'-PRIME,			
SOURCE	Drosophila melanogaster (Fruit fly)				KEYWORDS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
ORGANISM	Drosophila melanogaster				ORGANISM	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
DEFINITION	Drosophila melanogaster				REFERENCE	Hordeum vulgare subsp. vulgare			
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				AUTHORS	Euarchonta; Ecdysozoa; Bilateria; Animalia; Arthropoda; Insecta; Pterygota;			
VERSION	AA263299.2 GI:13766074				TITLE	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
EST.					JOURNAL	Beetles; Drosophilidae; Drosophila.			
KEYWORDS					COMMENT	1 (bases 1 to 779), Harvey D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S., and Rubin, G.M.			
ORGANISM	Drosophila melanogaster				COMMENT	On Jan 24, 1997 this sequence version replaced gi:1899358.			
DEFINITION	Drosophila melanogaster				COMMENT	Other ESTs: LD06393_3prime			
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT	Contact: stpleton, M.			
VERSION	AA263299				COMMENT	BRCGP			
KEYWORDS	Drosophila melanogaster				COMMENT	Lawrence Berkeley National Lab			
ORGANISM	Drosophila melanogaster				COMMENT	One Cyclotron Rd, Berkeley, CA 94720, USA			
DEFINITION	Drosophila melanogaster				COMMENT	Fax: 510 486 6798			
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT	Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu			
VERSION	AA263299				COMMENT	hit genomic AB003544: arm:3L [11370031,11651899]			
KEYWORDS	Drosophila melanogaster				COMMENT	estimated-cyto:68C3-68C15: 04/10/2001			
ORGANISM	Drosophila melanogaster				COMMENT	Plate: LD.63 row: H column: 9			
DEFINITION	Drosophila melanogaster				COMMENT	High quality sequence stop: 701.			
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT	Location/Qualifiers			
VERSION	AA263299				COMMENT	1. .779			
KEYWORDS	Drosophila melanogaster				COMMENT	/organism="Drosophila melanogaster"			
ORGANISM	Drosophila melanogaster				COMMENT	/mol_type="mRNA"			
DEFINITION	Drosophila melanogaster				COMMENT	/db_xref="BDGP EST:BDcln005720"			
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT	/db_xref="taxon:7227"			
VERSION	AA263299				COMMENT	Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.			
KEYWORDS	Drosophila melanogaster				COMMENT	Barley ESTs from germinating seeds			
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT				
ORGANISM	Drosophila melanogaster				COMMENT				
DEFINITION	Drosophila melanogaster				COMMENT				
ACCESSION	FBAN006084 GO: [enzyme (GO:0003824)] located On: 3J 68C9-68C9;				COMMENT				
VERSION	AA263299				COMMENT				
KEYWORDS	Drosophila melanogaster				COMMENT</				

Db	389	ACCCCTGGCGAGCTTGGCCTGACTTACCTTGACCTCTAACCTCATCCACTGCCCTGGC	448	QY	61	GGTACCTTCGCTTAGTGAAGGTCCAAAGGGAGAACCTTACATGCTTCAACCTGGCTG 120
QY	361	GCCGAGAAGATAAGGCGAGGTTGAGGCCAAGATTGGCCCTGAGGCCAAATAACGCTTCTTC	420	Db	209	GGACTT-----GGZAGGCCGCTCCGGCAAGGGCTGGGTCTTGGCTGCCTC 262
Db	449	TITGCTCCGGCCACCTTCCCCTCAAG---TCGAGAACGGTCCAGGTCAGCTC	505	Db	121	AAGACCGGTTACCGTACTTGGACTCTGCTGACTTACCTGAACTGGGTGAGTTGGT 180
QY	421	AAGGACTCTGACCGAGAACCGGACATGGGCCATGGGATTTATGAGGT	480	Db	263	AGGTGCTACCGCACCTGGCAAGGTACAGCAACAGAGGGTGGTGGT 322
Db	506	AA---CCAGAATGTTCCATGTCGAGACCTGGAGGCACTGACGGCTCAAGTC	562	QY	181	GAGGGTATCCGTAACCTGGACTTCCTGAGGGAAACCCCTCCGGTAAAGCTCTTCGTC 240
QY	481	CGCAAAGGCAAGTCCATTGGCATGGACATGGCCACTTGGAGAAAGATGTC	540	Db	323	GAGGCATCAAAGGGTCTTCGAGGCGAGGATATTTCATC 382
Db	563	--CAAAGTCCCTCCCTCCCTGGTCTCAATTACCATGAGACCCGGCCGATC	620	QY	241	TGCACCAAGGCTGGAAACCACCTCCACCGTTATGGAGCAGCNCCTCTGGTCCATGACGAC 300
QY	541	AAGTTGGCAAGGTCAAGGTCTGGCTCACGCCAACCTTACCATGAGATTCACCCCTCCGCCAAC	600	Db	383	ACCTCAAGCTGGTGGACAACCTCCAAAGCCGTTAGCTGGAAATAAGTCATTCCTC 442
Db	621	GAGGCCAACGGCGTGTCCGGCTAACAGATGAGGCAACCCCCCTCCGCCAAC	680	QY	301	TCCCTGAGCGCTTGGACTCTGACTGTTATGGAGGTTATGGAGATTTGGCATGTC 360
Qy	601	GAGGAGCTGGTGTGCAAGTCACTGTTCTCAAGACATT	638	Db	443	ACCCATGCCGAGCTGGCCTGACTTACCTCATCAGCTGCCCTGACCT 502
Db	681	CAGCCCCCTGATGCACTACTGGCCCAAGAACGGCAT	718	QY	361	GCGGAGAAGAATGCGCAGGGTGAAGCTGGCTGAGCCAAAGCTCCAAAGCTGGCTC 420
RESULT				Db	503	TTTGTCTGGGGGGGAGCTTCCCAAG--TCCGAGGACGGCTCCGGAGTGGCTGC 559
CB899137				QY	421	AAGGACCTGACCGAGACCCGACATGGCCGATATGGCTGAGGTTATGGAGATTTGGAT 480
LOCUS				Db	560	AA---CCAGAATGTTGCACTGCTCAAGCTGGAAAGCCATGACCGAGTGGCTGC 616
DEFINITION				QY	481	CGGAAGGCCAGGTCCATTGGTCTCAACTGGACATTGGGACCTTGAAGATGTC 540
trico17xh08				Db	617	--GAGGTCGCTCCGTCGCTCAACCTTACCATGGACCCGGACGCCCATC 674
T. reesei				QY	541	ANGTTGCGCAAGSTCATGGCTTACGGCAACCGATGGAGTTACGCCCTGCTGCCAAC 600
mRNA culture				Db	675	GAGGCACCCGGCTGCTGCTCCGGTAACCGATGAGGCACACCCGGCTCCCAAC 734
Version 3				QY	601	GAGGAGCTGGTGTGCAAGTCACTGCTCTCAAGACATAT 638
April				Db	735	CAGCCCCCTGATGCACTACTGGCCCAAGAACGGCATAT 772
Hypocreaceae				RESULT	12	
EST				AA785751		
KEYWORDS				LOCUS	AA785751	189 bp mRNA linear EST 31-JUL-1998
SOURCE				DEFINITION	AA785751	Aspergillus nidulans 24 hr asexual development and
ORGANISM				ACCESSION	AA785751	vegetative cDNA lambda zap library Emericella nidulans cDNA clone
Hypocreaceae; T. reesei; Hypocreales; Ascomycota; Sordariomycetes;				VERSION	AA785751	h1h0ai 5', mRNA sequence.
Fungi; Ascomycota; Pezizomycotina; Hypocreales; Hypocreaceae; Hypocreales.				KEYWORDS	AA785751.1	EST
REFERENCE				SOURCE		Emericella nidulans (anamorph: Aspergillus nidulans)
AUTHORS				COMMENT		Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella.
TITLE				REFERENCE	1	(bases 1 to 189)
JOURNAL				TITLE		Unpublished (1998)
COMMENT				ORIGIN		We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: SK High quality sequence stop: 182. Location/Qualifiers
Contact: Pamela K. Foreman				Query Match	16.4%	Score 160 4; DB 14; Length 823;
Genencor Intl.				Best Local Similarity	57.7%	Pred. No. 1.e-4;
925 Page Mill Road, Palo Alto, CA 94304, USA				Matches 368;	Conservative	O; Mismatches 256; Indels 14; Gaps 4;
Tel: (650) 846-7635				Matches 368;	Matched	
Fax: (650) 621-7817				/note=Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
Email: foreman@gencor.com				/dev stage="mycelia"		culture grown from 24 hrs to 6 days with varying Carbon
Seq primer: LR-F1 primer.				/clone="trico17xh08"		and Nitrogen sources and concentrations."
FEATURES				/mol type="mRNA"		ORIGIN
Source				/strain="OM6a"		Query Match
Location/Qualifiers				/db_xref="taxon:51453"		Best Local Similarity
1..823				/dev stage="mycelia"		57.7%; Pred. No. 1.e-4;
/organism="Hypocrease jecorina"				/clone="1..16"		Matches 368;
/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		Matched
culture grown from 24 hrs to 6 days with varying Carbon				culture grown from 24 hrs to 6 days with varying Carbon		Indels 14; Gaps 4;
and Nitrogen sources and concentrations."				and Nitrogen sources and concentrations."		
ORIGIN				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		
				culture grown from 24 hrs to 6 days with varying Carbon		
				and Nitrogen sources and concentrations."		
				/note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial		

Db	662	ATCAAGTCCAACTCCAGGTCTTGTATTCTTCAGGAGTTAGTTCTGGACAGC	721	Qy	241	TGCCACAGGTGTTGA	256
Qy	883	GTTGCCAAGG 892		Db	519	TGACCAAGGTCTGGA	534
Db	722	CTTGAGAGG 731					
RESULT 9				Qy	241	TGCCACAGGTGTTGA	256
AA746853	AA746853	534 bp mRNA linear EST 31-JUL-1998		Db	519	TGACCAAGGTCTGGA	534
LOCUS	mef07a1:rl1	Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone mef07a1 5', mRNA sequence.					
DEFINITION							
VERSION	AP746853	GI:2847183					
KEYWORDS	AP746853.1						
SOURCE							
ORGANISM	Emericella nidulans (anamorph: Aspergillus nidulans)						
ACCESSION							
EST.							
REFERENCE	AP746853						
AUTHORS	Kupper,D., Gray,J.J., Hausner,J.J., Lai,H., Martin,W., Aramayo,R.,						
TITLE							
JOURNAL							
COMMENT	Unpublished (1998)						
OTHER ESTS:	mef07a1:rl1						
CONTACT:	Bruce A. Roe, University of Oklahoma, broe@ou.edu						
Department of Chemistry and Biochemistry, University of Oklahoma							
Advanced Center for Genome Technology, University of Oklahoma							
620 Parrington Oval, Norman, OK 73019, USA							
Tel: 405 325 4912							
Fax: 405 325 7762							
Email: broe@ou.edu							
We anticipate the future release of the cDNA clones to the Fungal							
Genetics Stock Center							
Seq Primer: SK							
High Quality sequence stop: 522.							
Location/Qualifiers							
1..534	/organism="Emericella nidulans"						
/mol_type="mRNA"	/strain="FGSC A26"						
/db_xref="taxon:162425"	/db_xref="taxon:162425"						
/clone="IMCF01"	/clone="IMCF01"						
/tissue type="vegetative mycelia, asexual structures"	/tissue type="vegetative mycelia, asexual structures"						
/clone Lib="Aspergillus nidulans 24hr asexual	/clone Lib="Aspergillus nidulans 24hr asexual						
developmental and vegetative cDNA lambda zap library"	developmental and vegetative cDNA lambda zap library"						
/note="vector: pBluescript SK"; Site_1: EcoRI; Site_2:	/note="vector: pBluescript SK"; Site_1: EcoRI; Site_2:						
XbaI; 5' end of cDNA cloned into EcoRI site of pBluescript	XbaI; 5' end of cDNA cloned into EcoRI site of pBluescript						
3' end of cDNA cloned into XbaI site of pBluescript"	3' end of cDNA cloned into XbaI site of pBluescript"						
FEATURES							
source							
Query Match	16.9%	Score 164.8; DB 9;	Length 534;	Db	155	GGGGTACTCCGTGAGCTTCCGGGAGGGCTCGCTGTCCTGAGCCATCGCT	60
Best Local Similarity	77.7%	Pred. No. 9.7e-26;		Qy	95	ATGTCCTTCGGTACGGTCACTCCGGTACAGATCCGGCTCGCTGTCCTGAGCCATCGCT	154
Matches	199;	Conservative 0; Mismatches 57;	Indels 0; Gaps 0;	Db	61	GGTACCTTCGTTAGTCAAGTGTCCAGGGAGACTATACTGTCGTCCTG	120
Qy	1	ATGCTTAACGGAAGACTTCAATTGAGAACGGCAAGATTCCGGCGGGCTT	60	Qy	121	AGACGGTGTACCGTCACTTGACGTGCTGGTACTACTGACGAGGTGAGTTGGT	180
Db	279	ATGCTTAGCAGGAAAGACTTCAAGCTTCAAGAACGGGTACCATCCGGCGGGTTTC	338	Db	209	AGGGTTGGTACCGCCACCTGACCTGGCAAGGTCTACGGCAAACGGAGGTGGT	268
Qy	61	GGTACCTTCGTTAGTGAAGTTCAGGGAGACCTATCTGTACACATGCTG	120	Qy	181	GAGGGTACTCCGTTAGTGAAGGAGAACCTCCACGTTATGAGGAGCTTCGTC	240
Db	339	GGTACCTTCGTTAGTGAAGTTCAGGGAGACCTATCTGTACACATGCTG	398	Db	269	GAGGGATCAAGAAGCTTGTGAGTCCTGGCTGAGGCTCCGGCTGAAGCGGAGGATATTTCATC	328
Qy	121	AAGACGGTGTACCGTCACTTGACGAGGTGAGGTGGT	180	Qy	241	TGACCAAGGTGTGAAACCCTCCACGTTATGAGGAGCTTCGTCATGGCAACGAC	300
Db	399	GAGGTGGAGAACGGCACTGGACTGCTGGCTGTTCTACAGAACGAGGAGGTGGC	458	Db	329	ACCTCCAAAGCTTGTGAACTCCACAACTGGCGGCGTGTGAGAC	388
Qy	181	GAGGTGGATCCGTGACTTCTCTGAAGGAGAACCCCTGGTAAAGCTGAGACATCTTC	240	Qy	301	TCCCCCTGAGGTGTCTGGACTGTTGACTATGTTCTCGTCACTGGCCCATTTGT	360
Db	459	GATGCCGTTGCGACTTCTCTCAAGAGAACCCCTGTGTAAGCGGAGAACATTCATC	518				

Db	265	ACCGAAAGCTTGGAACACAAGCCCTTAAGGAACTGGCTCTTGACCA	324	Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactic, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
Qy	301	TCCCTGAAGGGCTGGACTGACTGATGGTCCTCGTCACTGCCATTGCT	360	
Db	325	ACCCCTGAGGCTGGCTGTGTTGTGTTGATCCACTGGCTGTGCTC	384	
Qy	361	GCGAGAAGATGGCAAGGTGAGCCAAATAGTCATTC	420	
Db	385	TTCAAGAACGGAAACGCTTCCACTGAGGAGATGGCTAGGAGCT	444	
Qy	421	AAGGACTGACGGAAACCCGACATGGGGCATATGGAAGATTATGAGGAT	480	
Db	445	GACCAAGGGTGGACCTCTCCAGACCTGTAACCAAGTGCCTAGGAG--	501	
Qy	481	CGCAAGGCAGGTGCTCCAACTGGGCAATTGGCTTCAAGGATGTC	540	
Db	502	--ANGTTGCTCCATTGCTGCTACTTCAAGAGATGCTGAAATCATC	558	
Qy	541	AAGTTCGCAAGGTGATGGGATGGGATCACCCCTTCCTGCCAAC	600	
Db	559	AAGGAACTGGCTGGGCTTCCAGGGGATGAAACAGATGGCTTCCNA	618	
Qy	601	GAGGACTGGTGGCACTACTSCTTCTCCAAAGAACATTATGCCGTCTCTG	660	
Db	619	CCTGAGCTCTCAAGTCACTTCAGGGATCTCTCTACTGCCMACTCGCCNTC	678	
Qy	661	GGCTGGCAGGAGCAG 676		
Db	679	GGCAACACAGCTGG 694		
RESULT	8			
LOCUS	CNS05Q60	1087 bp	DNA	GSS 05-TJL-2001
DEFINITION		linear		Yarrowia lipolytica, genomic survey sequence.
ACCESSION	AL10278			
VERSION	AL10278.1	GI:12178444		
KEYWORDS	GSS			
SOURCE	Yarrowia lipolytica			
ORGANISM	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccaromyctes; Diodontaceae; Yarrowia.			
REFERENCE	1	(bases 1 to 1087)		
AUTHORS	Souciet,J.L., Aigle,M., Artiguaneau,F., Blandin,G., Boletin-Fukuhara,M., Bon,E., Brattier,P., Casaregola,S., Del-Montigny,J., Duron,B., Llorente,B., Loubrie,A., Levinglie,A., Malenzy,J., Neuvilise,C., Olier-Kalogropoulos,O., Potier,B., Saurin,W., Tekaia,F., Toffano-Nioche,C., Wessolowski-Loulvel,M., Wincker,P. and Weissbach,J.			
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies			
JOURNAL	FEBS Lett.	487 (1), 3-12 (2000)		
MEDLINE	20584711			
PUBMED	11152876			
REFERENCE	2	(bases 1 to 1087)		
AUTHORS	Artiguaneau,F., Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feyrerol,C., Wincker,P. and Gaillardin,C.			
TITLE	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica			
JOURNAL	FEBS Lett.	487 (1), 95-100 (2000)		
MEDLINE	120584727			
PUBMED	11152892			
REFERENCE	3	(bases 1 to 1087)		
AUTHORS	Artiguaneau,F., Wincker,P. and Gaillardin,C.			
TITLE	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica			
JOURNAL	FEBS Lett.	487 (1), 95-100 (2000)		
MEDLINE	120584727			
PUBMED	11152892			
REFERENCE	4	(bases 1 to 1087)		
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-SEP-2000)			
MEDLINE	20584727			
PUBMED	11152892			
REFERENCE	5	(bases 1 to 1087)		
AUTHORS	Genoscope.			
TITLE	Submitted (07-SEP-2000)			
JOURNAL	Genoscope - Centre National de Sequencage, rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : secref@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,			
Qy	823	ATTGAACTTCAAACCTCAAGGCAATTGGCTCGATCAATGCC	882	

JOURNAL	for barley genomics: Fusarium infected Morex spike cDNA library
COMMENT	Unpublished (2001)
Contact	Wing RA
Clemson University Genomics Institute	
Clemson University	
100 Jordan Hall, Clemson, SC 29634, USA	
Tel: 864 656 7288	
Fax: 864 656 4293	
Email: rwing@clemson.edu	
Total hg bases	541
Seq primer:	AATTAACCTTCACTAAAGGG
High quality sequence stop:	785.
FEATURES	
source	<p>/organism="Hordeum vulgare subsp. vulgare"</p> <p>/mol_type="mRNA"</p> <p>/cultivar="Morex"</p> <p>/sub_species="vulgare"</p> <p>/db_xref="taxon:11509"</p> <p>/clone="FVSMB1009C04f"</p> <p>/tissue_type="spike"</p> <p>/lab_host="TJCL121"</p> <p>/clone_idb="Hordeum vulgare spike EST library HVCDNA0012 (Fusarium infected)"</p> <p>/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XbaI; Plants were grown at the University of Minnesota in the GJ Muehlbauer lab; spikes were harvested and snap frozen at 0, 1, 2, 3, 4, 5, 6, and 8 days after Fusarium graminearum inoculation (Heinen). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all eight RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Choi, Fenlon, Malatsasi). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmonds, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/projects/barley. Projects Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newslett 31:29-30. (http://wheat.pw.usda.gov/gggpages/bgn/31/cover.html)"</p>
ORIGIN	
Query Match	18.0%
Best Local Similarity	56.1%
Matches	379; Conservative 0; Mismatches 285; Indels 12; Gaps 60
Qy	1 ATGTCATAACGGAAAGACTTTACATTGAGCAAACGGGTCAAGATTCCTGGTGGCTTT
Db	31 ATGCTTTGCTGCGAAGCTGTACTCTAACCTGGTGGAGATCCCCAGATCGCTAC
Qy	61 CGTACCTTCGGTAGTGTAAGGTTCAAGGGCAGAACCTATCTGCCTGCTCACCACTGCTG
Db	91 GGCACAT-----GGCAGCTGCTCCGGTAAGGTGGCATGTTCTACAGGCCCTT
Qy	121 AAGACCGGTACGGTACCTGGTACTACCTGACACGGGTGAGTTGGT
Db	145 AAGGCTGGCTACCGTACCTGGTACTACCTGACACGGGTGAGTTGGT
Qy	181 GAGGGTATCCGTGACTCTCGTAAAGGAGAACCCCTCGTGAAGCTGAGGACATCTCGTC
Db	205 GAGGGTATCAAGAAGGCCCTCAAGGGGGTCTGGCTTAAGGAGCAGCATCTTCATC
Qv	241 TGGACCCAGGTGGAAACCACTTCAACGGTTATGAGGACCTCTGGTCAATGAGCAG

1.	.757	source	
/note="Vector: PREPY; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	/organism="Hypocrea jecorina"		
ORIGIN	/mol_type="mRNA"		
Query Match 22.6% Score 220.8, DB 14; Length 704;	/strain="Qka"		
Best Local Similarity 67.2%; Pred. No. 5.1e-18;	/dev_stage="mycelia"		
Matches 312; Conservative 0; Mismatches 152; Indels 0; Gaps 0;	/clone="tric005xa21"		
Qy 13 AACACTTACATTGAGAACCGGGTAAGATTCCMGGCCTGGTACCTTCGCT 72	/note="Vector: PREPY; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."		
Db 230 AACGGTCAACTCTGAACCCGTGCCAAGATACCCCGGTGGGTACATGCC 289	ORIGIN		
Qy 73 AGTGAAGGTTCCAAGGGGAGACCTATACTGCTGTCACACTGCCCAGAACCGGTTAC 132	Query Match 22.6%; Score 220.8; DB 14; Length 757;		
Db 290 AATGAGGGTGCRAAGGGGAGACATAGCACTGCTGTTAACAAAGCACTGAC 349	Best Local Similarity 67.2%; Pred. No. 5.3e-38;		
Qy 133 CGTCACTTGGACTGTGCTGTGCTGACTACTGACGAGCTGGTGGAGGTATCCGT 192	Matches 312; Conservative 0; Mismatches 152; Indels 0; Gaps 0;		
Db 350 CGCACCTGTGATTCGGCTGGTTAACACACAGAATGAGGTGTGAGC 409	Qy 13 AAGACTTACATTGAGAACCGGGTCAAGATTCGCTGGTACCTTCGCTTGCT 72		
Qy 193 GACTTCTCTGAAGGAGAACCCCTGGTCAAGCCTGAGGACATCTCTGACCAAGGTG 252	Db 283 AAGGCTGACTACTCTGAACCCGTGCCAGAACGAGCTGGTGGGGTGGGGTACATGCC 342		
Db 410 GATTTTCCTGCCGGCCGACGCTTAACCCGGAGATCTCTCATTTGACCAAAGTT 469	Qy 73 AGTGAAGGTTCCAAGGGGAGACCTATACTGCTGTCACACTGCCCAGAACCGGTTAC 132		
Qy 253 TGAACACCCTCCACCGTTATAGGAGCTCTCTGGTCATTGACACACTCCCTGAGCGT 312	Db 343 AATAGGGGCCCCAGGGAGACATCCAGCTGTTACACAAAGAATGAGGTGTGACCTGATGCC 402		
Db 470 TGAAACACCCTGTGATGAGCCAGGGCTCAAGTGGAGCCAAAGACTCTGCSAAAAC 529	Qy 193 GACTTCTCTGAAGGAGAACCCCTGGTCAAGCCTGAGGACATCTCTGACCAAGGTG 252		
Qy 313 CTGGAGCTTGACTACGTTGATAATGTTCTCGTGCCTACTGCGCCATTCTGGCAGAGAAAT 372	Db 403 CGCCACCTGTGATTCGGCTGGCTGTGTTACACAAAGAATGAGGTGTGACCTGATGCC 342		
Db 530 CTCAGGTGATCATGACTGCTGTTCTCGTGCCTACTGCGCCATTCTGGCAGAGAAC 589	Qy 193 GACTTCTCTGAAGGAGAACCCCTGGTCAAGCCTGAGGACATCTCTGACCAAGGTG 252		
Qy 373 GGCCAGGGTGAGGCCAAAGATTGGCCCTGAGGGCAAATACTGCTCATTCAGGACCTGACC 432	Db 463 GATTTTCCTGCCGGCCGACGCTGAAACCGGAGATCTCTCATTTGACCAAAGTT 522		
Db 590 AGGCACAGGAGCTCAACTGGCCGCTATGGGAGATGTTGATCACAGGCCCTGAGC 649	Qy 253 TGGAACTACCTCCACCGTTATGAGGACTCTCTGGTGTCACTGACGACTCTCTGGTGAAGCGT 312		
Qy 433 GAGAACCCCGAGGCCAACATGGGGCTPATGGAGAAGATTATGAA 476	Db 523 TGGAAACACCCTGTGATGAGCCAGGGACTCTGAGGACACTGAGCTGAACTGCTGCGGAAAC 582		
Db 650 GATAAACCCAGAGCCAACATGGCCAGGGCTATGGATGAGCTTGTTGA 693	Qy 313 CTTGGACTTGACTACGTTGATAATGTTCTCGTGCCTACTGCGCCATTCTGGCAGAGAAAT 372		
RESULT 5 CB896064 757 bp mRNA linear EST 02-JUL-2003	Db 583 CTCAGGTGATCATGACTGACCTGCTCTGGCCCATCTGGCAAGAATGGTCACTGCC 582		
DEFINITION tric005xa21 T. reesei mycelial culture, Version 3 april Hypocrea jecorina cDNA clone tric005xa21, mRNA sequence.	Qy 433 GAGAACCCGGAGGCCAACATGGGGCTPATGGAGAAGATTATGAA 476		
ACCESSION CB896064.1	Db 703 GATAAACCCAGAGCCAACATGGCCAGGGCTATGGATGAGCTTGTTGA 746		
VERSION EST.			
KEYWORDS Hypocreomycetidae; Hypocreales; Hypocreaceae; Sordariomycetes; Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; 1 (bases 1 to 757)			
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei) EST.			
ORGANISM Trichoderma reesei			
REFERENCE Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Dierer, S., Dunn-Coleman, N.S., Goeddegebuur, F., Houfek, T.D., England, G.J., Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchison, C., Olivares, H.A., Teunissen, P.J., Yao, J., and Ward, M.			
AUTHORS Transcriptional regulation of biomass degrading enzymes in the filamentous fungus Trichoderma reesei			
JOURNAL J. Biol. Chem. 278 (34), 31988-31997 (2003)			
PUBLISHED 22803114			
COMMENT Contact: Pamela K. Foreman			
GENERIC INTL. 925 Page Mill Road, Palo Alto, CA 94304, USA			
MEDLINE			
SEQUENCE PRIMER 1788890			
COMMENT Seq Primer: LT-F1 primer.			
LOCATION/Organizers Location/Qualifiers			
FEATURES Unpublished (2001)			
COMMENT Other ESTs: n210f5..f1			
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu			
COMMENT Department of Chemistry and Biochemistry			
COMMENT Advanced Center for Genome Technology, University of Oklahoma			
COMMENT 620 Parrington Oval, Norman, OK 73019, USA			

Qy	73	AGTGAAGGTTCCAAGGGGAGACCTATACTGCGTCACCAACTGCCTGAAGAACCGGTATC
Db	281	AATGAGGGTGCAAGGGAGACATAGGGCTTAAAGGACTGAGTTGGAAC
Qy	133	CGTCACCTGGACTGTGCTGGTACTACCTGAAACGGGTGAGTTGGGTATCCGT
Db	341	CGCCACCTGGTGGCTGGTACACAAACGAGATGAGTTGCGGTGAGCGGTAGCG
Qy	192	GACTTCCTGAGGAAACCCCTCGTGAACGGCATCTGGCACAGAGT
Db	401	GATTITCTGCCCGCGACCCGAGCTGACGGAGATCTCTCATGCAAAAGT
Qy	252	GACTGAGCTGGTCACTGGCCATCTGGGAGGATTTAGGATGCGAAGGCAAG
Db	460	TGGAAACCCCTCCAAGGTTATGAGGAGCTCTCTGGCCATGGAGAAAT
Qy	312	TGGAACCCCTCCAAGGTTATGAGGAGCTCTCTGGCCATGGAGAAAC
Db	461	TGGAACCCCTGAGGCAAGTGGCCAGGAGCTGAGGGCAAGTGTGAAAC
Qy	372	CCTGGACTTGACTAGTGTGATATGTCTCGTCCATTGCGAGAAAT
Db	521	CTCAAGGTCATTACATGACCTGTCCTGCACTGGCAATGCGCCAGAAC
Qy	580	CGGCCAGGTGAGGCCAATAGTCATTCTGAAGGACTCTGGGCC
Db	640	AGCGAGGGGGGCGTGAAGTGGCGGATGTCATCAAAGGCTGAG
Qy	492	GAAACCCCGAGCCACATGGCGCTTGGAGAAGATTATGGAGTCAGG
Db	581	GAAACCCAGGCCACATGGCGCTTGGAGAAGATTATGGAGTCAGG
Qy	493	TCCATTGGTCTCCAACTGGCACATTGGGAGTGTGCGATGTCAGGT
Db	643	TGGAAACCCCTGAGCTTCAATGGGAGTGTGCGATGTCAGG
Qy	584	GCAATGGGACTATCCAACCTGGACATNNGGGGTGAAGAGTCCT
Db	702	ATCAAGCCGGCAGTGAACCAAGGATGAGATCA
Qy	734	ATCAAGCCGGCAGTGAACCAAGGATGAGATCA
Db	792	ATCAAGCCGGCAGTGAACCAAGGATGAGATCA
RESULT	3	
LOCUS	CB897574	CEB897574 792 bp mRNA linear EST 02-JUL-2003
DEFINITION	tric01lxm03	T. reesei mycelial culture, version 3 april Hypocreae
ACCESSION	CB897574	jecorina cDNA clone tric01lxm03, mRNA sequence.
VERSION	CB897574.1	
KEYWORDS	EST.	
ORGANISM	Hypocreales	Hypocreales; Hypocreaceae; Hypocreales; Hypocrea; jecorina (anamorph: Trichoderma reesei)
REFERENCE	1	(bases 1 to 792) Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedgebuer, F.J., Houfek, T.D., Hutchinson, C., Kelley, A.S., Meeman, H.J., Mitchell, T.D., Olivares, H.A., Reunissen, P.J., Yao, J. and Ward, M. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei. J. Biol. Chem. 278 (34), 31988-31997 (2003)
AUTHORS		
COMMENT	Pamela K. Foreman	Contact: Pamela K. Foreman
JOURNAL	Genencon Int'l.	925 Page Mill Road, Palo Alto, CA 94304, USA
MEDLINE	22803314	Pubmed ID: 12788920
FEATURES	source	Organism: "Hypocreales"
		/organism="Hypocreales"
		/mol_type="mRNA"
		/strain="OM45"
		/sub_xref="taxon:51453"
		/clone="tric01lxm03"
		/dev_stage="mycelia"
		/clone_lib="T. reesei mycelial culture, Version 3 april"
		/note="vector: pREP3Y; Site1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
ORIGIN		
Query Match	27.1%	Score 265; DB 14; Length 792;
Best Local Similarity	66.3%	Pred No. 9 ge-48;
Matches	379;	Conservative 0; Mismatches 193; Indels 0; Gaps 0;
FEATURES	source	
		Organism: "Hypocreales"
		/organism="Hypocreales"
		/mol_type="mRNA"
		/strain="OM45"
		/db_xref="taxon:51453"
		/clone="tric01lxm03"
		/dev_stage="mycelia"
		/clone_lib="T. reesei mycelial culture, Version 6 October
FEATURES	source	
		Organism: "Hypocreales"
		/organism="Hypocreales"
		/mol_type="mRNA"
		/strain="OM45"
		/db_xref="taxon:51453"
		/clone="tric01lxm03"
		/dev_stage="mycelia"
		/clone_lib="T. reesei mycelial culture, Version 6 October

Email: r.bruskiwicz@cgiar.org
 International Rice Information System (IRIS);
<http://www.iris.irri.org/>

Assignment of putative function to the sequence by S. Rudd of the
 Munich Information Center for Protein Sequences
 (<http://mips.gsf.de>)

Plate: 03 row: B column: 08.

FEATURES

source

1. .886
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="IR64"
 /db_xref="taxon:39946"
 /cloner="FC000800"
 /issue_type="Panicles"
 /dev_stage="Flowering"
 /clone_lib="IRRI Drought Stress Panicle Library"
 /note="Vector: pBluscript II SK; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

ORIGIN

Query Match 36.1%; Score 353.2; DB 14; Length 886;
 Best Local Similarity 68.2%; Pred. No. 3.8e-67;
 Matches 527; Conservative 0; Nsmatches 242; Indels 4; Gaps 3;
 Qy 5 CTAACGGAAAGACTTCACATTGACCAAGCGGTAAAGATTCTGGCGTGGCTTTGGA 64
 Db 114 CAAACAGAAACATAAACAGTCAACATGGCTAACATGGGTAGGATTGGAA 173
 Qy 65 CCTTCGTTAGTAGTGAAGTTCCAAAGGGAGAACCTATAGTGTGACCAACTGCTGTGAA 124
 Db 174 CATTCGCAAAAGGGCTCCTAGGGGAACCTAGCTGCTGAACTGTGCTGAA 233
 Qy 125 CGGGTTACCGTCACTTGAACTGCTGGCTGACTACTGTGAGGAGGTGGTGGAGG 184
 Db 234 CTGGTACCGACACTCTGCACTGCACTGGTCACTGGTGGAGGGTGGAGG 293
 Qy 185 GTATCGGTGACTTCTCTGAAAGGAGAACCCCTGGGTGAAGAACCTTGTGCGA 244
 Db 294 CGTCGGGACTTCTGCCAGAACTCCCTGGTACTCTCCAAGAACATCTCTGCA 353
 Qy 245 CCAAGGTGAAACCACCTCCACCGTTAACGGAGCTCTCTGGTCAATGAGCACTCC 304
 Db 354 CAAAGGTCTGAAACCACCTCCACCGTTAACGGAGCTCTCTGGTCAAGCTGC 413
 Qy 305 TGAAGGCTCTGGACTACTGTGATAGTGTGTTCTGGTCACTGGCATGTCGC 364
 Db 414 TCGGAAAGTTCAGGGACTACATGATCTCTCTGGTCACTGGTCTGCTGTCG 473
 Qy 365 AGAAAGATGGCAGGGTGAAGCTGACGCCAAATGGTCACTTCAGG 424
 Db 474 AGAAAGACGAGAAATCATGCCAAAGCTTGAGCTGATGAGTACATCTGAGG 533
 Qy 425 ACCTGACGGAAACCCGACCTGGGCTATGGAAAGATTATGGGATCCGA 484
 Db 534 AGCTGACGGAAAGCTGCTCAACTGGGACCTGGGCAATCTGAACTCTGGA 593
 Qy 485 AGGCCAGGTCAATTGGTGTCTCAACTGGGACCTGGGCAATCTGAACTCTGGA 544
 Db 594 GGGCTAGGTATGGTCAACTGGGCAATCTGAACTCTGAACTCTGCACTGGT 653
 Qy 545 TGGCCAAAGGTCAATTGGTGTCTCAACTGGGACCTGGGCAATCTGAACTCTGGA 603
 Db 654 TCGGAAAGGTCAAGNCGGTCAACCTGGTCAACTGGGCAATCTGAACTCTGGA 713
 Qy 604 GAGCTGTTGAGTACTGTTCTCAAGAACATATGGCCGTACTCTGGTCACTGGC 663
 Db 714 GAGCTGTCAGTGTGAGTACTGGGCAATCTGAACTCTGGTCACTGGTCACTGGT 773
 Qy 664 TCGCAGAACGGTCCACACGGTCAAGGGTCAAGGAGAAACAGACTCTGAA 723

Db 774 TCCANAAANCAGTCCCA-CACAGCGAGGGTCTGAACTTCCAGTC- AACGAG 830
 Qy 724 ATGCGCAGAAGGGGGCAACACCTTGTGCTAGGTTCTATGCTGGGGCT 776
 Db 831 GTTGCTGAGCGCANCGGCAAGATCTGCAAGTCTGCAAGTCTGGA 883
 RESULT 2
 CF867536 LOCUS tric01lxm03.b1 T.reesei mRNA EST 31-OCT-2003
 DEFINITION tric01lxm03 mRNA culture, Version 6 October 2003
 ACCESSION CF867536
 VERSION CF867536.1 GI:38122162
 KEYWORDS EST.
 SOURCE Hypocrease jecorina (anamorph: Trichoderma reesei)
 ORGANISM Hypocreales: Pezizomycota; Sordariomycetes;
 Eukaryota; Fungi; Ascomycota; Pezizomyces; Hypocreaceae; Hypocreales; Hypocreaceae; Hypocreales.
 REFERENCE 1 (bases 1 to 734)
 AUTHORS Diermer,S.E., Dankmeyer,L., Dunn-Coleman,N., Houfek,T.D.,
 Mitchell,T.K., van Solingen,P., Teunissen,P.J.M., Ward,M. and
 Dean,R.A.
 COMMENT Unpublished (2003)
 CONTACT Ralph A. Dean
 Fungal Genomics Laboratory
 North Carolina State University
 Campus Box 7251 Raleigh, NC 27695, USA
 Tel: 919-513-0020
 Fax: 919-513-0024
 Email: ralph.dean@ncsu.edu
 Seq primer: IT-F1 primer
 Location,Qualifiers 1..734
 /organism="Hypocreale jecorina"
 /strain="OM6a"
 /mol_type="mRNA"
 /db_xref="taxon:51453"
 /clone="tric01lxm03"
 /dev_stage="mycelia"
 /clone_lib="T.reesei mycelial culture, 2003"
 /note="Vector: PRE3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
 ORIGIN
 Query Match 27.1%; Score 265; DB 14; Length 734;
 Best Local Similarity 66.3%; Pred. No. 9.5e-48;
 Matches 379; Conservative 0; Mismatches 193; Indels 0; Caps 0;
 Qy 13 AGAACATTCACTTGGCAACGGCTCAAGTTCTGGTACCTTGCT 72
 Db 163 AACAGCTAACCTGAAACCCGGTCAAGATCTGGCAATCTGCCC 222
 Qy 73 ATGGAAGTCAAGGGCTGAGGCTTACCTGCTGAACTGCTGAA 132
 Db 223 AATGAGGCTGCAAGGGCAGACATGCACTGGCTGGTGGATAC 282
 Qy 133 CGTCACATTGGACTGGCTGGTACTACTGAACTGGTGGATAC 192
 Db 283 CGGCACCTTGTGCTGGCTGGTGGATGGTGGATGGTGGATGG 342
 Qy 193 GACTTCTGAAAGGAGAACCCCTGGTGAAGCTGACATCTGACAGGT 252
 Db 343 GATTTCCTGGCGCCACCCACGGTCAAGGAGATCTGGTGGATGG 402
 Qy 253 TGGAAACCCACCTGGTCAAGGGTCACTGGTGGATGGTGGATGG 312
 Db 403 TGGAACCCACCTGGTCAAGGGTCACTGGTGGATGGTGGATGG 462
 Qy

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2004, 11:17:03 ; Search time 2523 Seconds

(without alignments)
11575.590 Million cell updates/sec

Title: US-10-004-115B-2

Perfect score: 978

Sequence: 1 atgtcttaaaggaaaggacttt.....ccaaagacatgttcgtcgatgaa 978

Scoring table: IDENTITY NUC

Gapov 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
Maximum Match 100%
Listing First 45 summaries

Database : EST:*

1: em_estba:
2: em_estbam:
3: em_estin:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_htc:
9: gb_sst1:
10: gb_est2:
11: gb_htc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estion:
16: em_eston:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_rnt:
21: em_gss_sun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_fod:
26: em_gss_phg:
27: em_gss_irrl:
28: gb_gss1:
29: gb_gss2:
30:

RESULTS

RESULT 1
CA764322

LOCUS AF55-RPf_03_B08_T7_018_abl

DEFINITION Oriza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

ACCESSION CA764322

VERSION 2

EST.

SOURCE Oriza sativa (indica cultivar-group)

ORGANISM Oryza sativa (indica cultivar-group)

COMMENT Unpublished (2002)

On Dec 2, 2002 This sequence version replaced gi:25993577.

REFERENCE Bennett,J., Arumugam,K., Lafitte,R., Wen,J., Rudd,S. and Bruskiewich,R.M.

TITLE IRRI Drought Stress Panicle cDNA Library

JOURNAL Unpublished (2002)

COMMENT On Dec 2, 2002 This sequence version replaced gi:25993577.

REFERENCE Bennett,J., Arumugam,K., Lafitte,R., Wen,J., Rudd,S. and Bruskiewich,R.M.

TITLE IRRI Drought Stress Panicle cDNA Library

JOURNAL Unpublished (2002)

COMMENT On Dec 2, 2002 This sequence version replaced gi:25993577.

REFERENCE Bennett,J., Arumugam,K., Lafitte,R., Wen,J., Rudd,S. and Bruskiewich,R.M.

TITLE International Rice Research Institute

JOURNAL Biometrics and Bioinformatics Unit

COMMENT International Rice Research Institute

Address: Metro Manila, Philippines

Phone: +63-2-845-0563

Fax: +63-2-845-0606

Result No.	Score	Query Length	DB ID	Description
1	353.2	36.1	CA764322	CB764322 AFP3-RPf
2	265	27.1	734	CB86536 tric011xm
3	265	27.1	792	CB897574 tric011xm
4	220.8	22.6	704	CF866105 tric005xa

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	353.2	36.1	CA764322	CB764322 AFP3-RPf
2	265	27.1	734	CB86536 tric011xm
3	265	27.1	792	CB897574 tric011xm
4	220.8	22.6	704	CF866105 tric005xa

Qy 623 TCTCCAGAACATTATGCCGGCTACTCTCCTCTGGCTCCAGA---CGGGTTC 679
 Db 668 AGTCAAAGCACATTACAATTAACGCTAAAGTCCTGGATCAGGCCCATTGG 727
 Qy 680 CCACCACTGGTGGGGTCACTGAGAAGACTCTGAACGAGATGCCGAGAG 736
 Db 728 CCAAGCTGTGTGATCCGTATCCCTAGGGCTTAAGGTCAGGAANTGCCCTAAGA 787
 Qy 737 GGGCAACACCTTGTCAGTTCTTATGCCCTGGCTGCCGTACCTCGTTC 796
 Db 788 AGAAGAGACCCCTGGACAGATCCTTATGCCCTGGCTGCCGTACCTGTTC 847
 Qy 797 TCCCCAAGAGCTCCACCCCAAGGCCATGAGTCAACTCAAGAGCATTGAGCTC 852
 Db 848 TCCCCAAATCTGTGACCAAGACCGATCACTCCAGTCAACTCCAGTCTGACTC 903

RESULT 15
 ID ABZ56044 standard; cDNA; 403 BP.
 XX DT 28-MAR-2003 (first entry)
 DE Aspergillus oryzae polynucleotide SEQ ID NO 5157.
 KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
 KW expressed sequence tag; gene; ss.
 OS Aspergillus oryzae.
 XX PN WO200279476-A1.
 XX PD 10-OCT-2002.
 XX PF 22-MAR-2002; 2002WO-IB000890.
 XX PR 30-MAR-2001; 2001JP-00098371.
 PA (NAAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (NARE-) NAT RES INST BREWING.
 PA (NORQ) NAT FOOD RES INST MIN AGRIC.
 XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
 Takeichi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
 DR WPI; 2003-046817/04.
 XX PT Detection of expression of specific Aspergillus genes for monitoring the
 PT fermentation and growth conditions of the fungus, using DNA probes.
 XX PS Claim 1; SEQ ID NO 5157; 48pp + Sequence Listing; Japanese.

CC The invention relates to polynucleotide having any of 6006 specific
 CC specific culture conditions, including one or more of eutrophic,
 CC oligotrophic, solid, early germination, alkaline, high temperature, low
 CC temperature or maltose culture or poly nucleotides stringently hybridising
 CC to these sequences. The polynucleotides are useful for monitoring the
 CC progress of fermentation and the growth conditions of a fungus,
 CC especially of Aspergillus oryzae which is widely used in industrial
 CC fermentation. Also monitoring for fungal contamination. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)
 XX SQ Sequence 403 BP; 103 A; 107 C; 101 G; 92 T; 0 U; 0 Other;
 XX Query Match 18.0%; Score 176.4; DB 7; Length 403;
 Best Local Similarity 70.9%; Pred. No. 3e-37;
 Matches 248; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

Search completed: May 29, 2004, 11:11:56
 Job time : 455 secs

Qy 1 ATGCTAACGGAAAGACTTACATTGAGCAACGGGTCAAGATCTGGGCTCAGCTT 60
 Db 54 ATGCTCTCTGGCTAACCTGAGCTAACGGGTCAAGCTGGCTCCCT 113
 Qy 61 GGTACCTTCGCTAGTGAAGGTTCCAGGGAGACCTATACTGCCTGCTAACCTGCTCCCT 120
 Db 114 GGTTACCTTCGCTAACAGGTCTCCAGGAGACTAACAGGTCTCCGTACGGGTTG 173
 Qy 121 AAGACCGGTTACGGTACTTGGACTGTCAGTGGTACTACCTGAACGAGGTGAGGTGGT 180
 Db 174 AAGTCGGTACGGCTACGGCACTCTGAATGCTGGTTATGCCAAAGGAGGAGTAGGA 233
 Qy 181 GAGGGTATCCGGTACTCTGGTGAAGGACATCTCGTC 240
 Db 234 CAGGGATTAAGATTCCTGAAGAAGGAAACCGTGTAAACGGAGGATCTGTTGTG 293
 Qy 241 TGCACCAAGGTGGAAACCACTCCACCGTATGAGGAAACCCCTGGTGAAGCCATT-GACGA 299
 Db 294 ACCACAGGGTGGAAACCACTGGTGAACCTGGTGTGGTGGAAAGA 353
 Qy 300 CRCCCTTAAGGCTCTGGACTTGACTTGACTGACTGTTCTCGPTCACT 349
 Db 354 TTCAATGAAAGAACTTGGAAATTGAGCTACGGGATCTGTTCTCGTCACT 403

Db 104 GGGCAGGTTGGTCCAGGGCTCAGGTCCCATCGAGTCAGTGGATACCTCGACCT 163
 Qy 147 TGCCTGGTACTACCTGAACAGGTGAGCTTGGTGGGAGGTTCTGACTCTGGAGGA 206
 Db 164 TGCAAGGTTACTCGAACAACTGAGTGGTGGGCCATA-----A 208
 Qy 207 GAAACCCTGGTGAAGGGTGGGAGGACATCTGACCAAGTGCGAACCTCGCA 266
 Db 209 GGAGCTGGCTAGCGCAAGGACCTCTGAGAACACTCGAACACTCGCA 268
 Qy 267 CGGTATGAGGAGCTCCTGGTCAATTGAGGACTCCCTGAAAGCTTGGACTTGACTA 326
 Db 269 CGCCGGGAGCAGGGCTGGCCCTTGAGAACCTCAAGAGCTGGCTCGAGTA 328
 Qy 327 CTGTTATGTTCCCTGTTGTTGCACTGGCCATTGCTCGAGAAATGGCCAGGGTGAGCC 386
 Db 329 CCTCCACCTTAACTCATICACTGGCCCTGGTTCGGCGAGGCGACATCACCCA 388
 Qy 387 CAAGATTGGCCATGAGGAAATACTGATTCATCTGAAG--GACCTGACCGAGAACCCGA 443
 Db 389 GAACCTCTCCGGAGGGCAGCAGAACGGTCAAGGTCAGCTGGAGTCAGCTGGCTGT 448
 Qy 444 GCCCCATGGCGCTATGGAGGAGATTATGGAGATGCCAAGGCCAGGTCCATTGGGT 503
 Db 449 CGAAGCTGGAAAGGGATGGTCAGCTTCAGCTTCAGCTTCAGCTCAAGGGATCGAGGT 508
 Qy 504 CTCCCAACTGGACCATGGCCGACCTTGGCAAGGGTGGCAAGGTCAATGCCCTCA 563
 Db 509 TTCCCAACTTGGCAAGATGGTGGTCAACCGCCATCATCGAGGTTACCGCGTGAACCCCTC 568
 Qy 564 CGCCACCACATGGAGATTACCCCTCTGCCAACGAGACTGGTGGTAGTACTCTT 623
 Db 569 CGTCACCAAGATCGAGTCACTCCCTCTCCAGGCCACCAAA 628
 Qy 624 CTCAGAAATTATGCCCTGGGCTACTCTCCCTGGCAACAGGTTCCOAC 683
 Db 629 GGCCAGAACATTACATACCGGATACTCTCTCGTAAACACCGTGGGGGCC 688
 Qy 684 CACCGGTGACGGGTCTGGGCTACTCTCCCTGGCAACAGGATGCCAGAAGGGCGCAA 743
 Db 689 TCTTGTGTCAGACCCGGAGATAAGGC-----ATGCCAGAGAACGCGCTG 739
 Qy 744 CACCTTGTCTGAGGTCTTATTGGCTGGGCTCTGGCTGCGCTGGCTACGTTCTCCCAA 803
 Db 740 CACGCCGCTGAGGTCTGGGCTACTCTGGGCTACTGGGCTACCTGGGTTATCCCAA 799
 Qy 804 GAGCTCCAACCCAAAGCGCATTTGAGTCCAACTCTCAAGAGCATGAGCTCTCGATGCCA 863
 Db 800 GTCGTCAACCCCTCCGGATTTGGAGAACCTTCAAGGGTCTCCAGGAGGA 859
 Qy 864 CTTTCAAGGCATCAATGCCCTGGCAAGGGTGTCACTTCGTTGTCACAT 917
 Db 860 CGTCAATGCCCTGGCAAGGCCATGGGCTGGCTAACACAT 913
 Db 380 AGTACCTCGTACTGACTATCCACTGGCCATGGTACAAGAGG-----G 427
 RESLT 14
 ABL0491 ID ABL0491 standard; CDNA; 978 BP.
 XX AC ABL0491:
 DT 26-MAR-2002 (first entry)
 XX XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 25955.
 DE XX Drosophila; developmental biology; cell signalling; insecticide;
 KW XX pharmaceutical; gene; ss.
 OS XX Drosophila melanogaster.
 PN XX WO200171042-A2.
 XX 563 AGCCCAACAGATGGGATTCACCCCTCTGGCCAAAGGACTGGGAGACTGT 622
 DB 608 TAACCCATCAGTGGCCTCATCATCTGACCTGGCAAGAGGAGCTGCTGCA 667

PD XX 27-SEP-2001.
 PR XX 23-MAR-2001; 2001WO-US009231.
 PR XX 23-MAR-2000; 2000US-0191637P.
 PR XX 11-JUL-2000; 2000US-00614150.
 PA (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW,
 PR XX WPI; 2001-656860/75.
 DR XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
 XX PS Claim 1; SEQ ID NO 25955; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (AB157737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences
 SQ Sequence 978 BP; 277 A; 241 C; 242 G; 218 T; 0 U; 0 Other;
 Query Match 18.3%; Score 178.8; DB 4; Length 978;
 Best Local Similarity 55.8%; Pred. No. 9.5e-38;
 Matches 433; Conservative 0; Mismatches 322; Indels 21; Gaps 4;
 Qy 83 CCAAGGGGAGACCTATACTACTGCTGTCACACTGCCTGAAAGACCGTTACCGTCACITCG 142
 Db 143 CCAAGGGTCAAGGTCAAGGGCTGCAAAACTGGCATATGATGGGGATACAGGACATG 202
 Qy 143 ACTGGCTCTGGTACTACCTGAAACGGGGTGGGTTGGTGGGTTACCTGGTACTCTCTG 202
 Db 203 ACTGGCTCTATGTATCACAAACGGGATGAGGGTGGTGT-----GAGGCGCAAGA 259
 Qy 203 AGGAGAACCCCTGGAAACGGTGGGCTCTGGTCTGGCAACAGGTTGGGACACCC 262
 Db 260 TCAAGGGGGGGTCAAGGGGGTGGATCTGGTCACTACCGGAAACCTGGGAAACACTT 319
 Qy 263 TCCACCGTTATGAGGAGCTCTCTGGTCAATGAGCTGGCTCTGGACTCTGGACTG 322
 Db 320 TCCATGCCCTGGATCTGGTCACTGGCCATGTCGGAGAGAACATGGCTCTGGTCA 379
 Qy 323 ACTACCTTGATATGTTCTCGTGTCACTGGCCATGTCGGAGAGAACATGGCTCTGG 382
 Db 380 AGTACCTCGTACTGACTATCCACTGGCCATGGTACAAGAGG-----G 427
 Qy 383 AGCCCAAGATGGGCAATACTCTCATGGTACGGAGACCCCG 442
 Db 428 GATGCCATCTTCCACCGACAGGATGAGCTGGGGTGTACTCGGGTGTATGATAG 487
 Qy 443 AGCCCAQATGGGGCTATGGAGAATTTATGAGGATGGCAAGGCCAGTCATTGGTG 502
 Db 488 TCGACAGTGGTCAAGGCTATGGGAGAGTGTGGGAGAGTGTGGCTATGGTGG 547
 Qy 503 TCTCCAACTGACCATGGCATCTGGGACCTTGTCAAGATGTCAGTTGGCAAGGTCTGGCTC 562
 Db 548 TTTCGAACTICAACAGAGGAGATGAGGGCTGGCTGCTGGTGGCCACTATTCCACZAG 607
 Qy 563 AGCCCAACAGATGGGATTCACCCCTCTGGCCAAAGGACTGGGAGACTGT 622
 DB 608 TAACCCATCAGTGGCCTCATCATCTGACCTGGCAAGAGGAGCTGCTGCA 667

Designing synthetic nucleic acid sequences for improved amplification, expression in host cell, by comparing free energy of folding of a starting polynucleotide and a modified polynucleotide having a codon replacement.

Example 9; Page 62-65; 117pp; English.

The present sequence encodes a NADPH-dependent aldehyde reductase (ALR). The polynucleotide sequence was modified using using the method of the invention. The specific invention describes a method for designing a synthetic polynucleotide. The method comprises providing a starting polynucleotide, determining the predicted free energy of folding per base of the polynucleotide, modifying the polynucleotide by replacing a codon with a different codon to provide a modified polynucleotide, determining free energy of folding per base of the modified polynucleotides, and comparing this with that of the original polynucleotide. The method is useful for developing nucleic acid sequences that enhance expression of the encoded protein in a heterologous host. The design and preparation of the synthetic genes are used in application of gene shuffling, directed evolution and molecular breeding methods. The method allows expression of genes from various organisms such as mammals, plants, yeast, fungi and bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic hosts at commercially viable levels, in particular proteins with low yield such as methionine gamma-lyase from *P. Putida*. (Updated on 11-SEP-2003 to standarise OS field)

Qy	684	CACCGTGTAGGGTCAAGAACAGACTCTAACGCGATCCGAGAACTGATCCGAGATCCGAGAAGGGGCAA	743
Db	660	TCTTCCTTCACCAACCGGAGATAAGGC-----	710
Qy	744	CACCTTGCTCAAGTCATTATGCCATGGCTGCCATGGCTGCGCTGGTAATGTTCTCCCAA	803
Db	711	CACGGCCGCTCAGGCTCTCATGGCTGCCATGGCTGCGCTGGTAATGTTCTCCCAA	770
Qy	804	GAGGTCAACCCAAAGCGATTAGTCGAACCTGAAGCACATTGCTCCGATGCTTCGAGGAGCTCAAGGAGCTCAAGGAGCTTCGCTCTGCAGGAGGA	863
Db	771	GTCGTCAACCCCTCCCGATTGCGAGCTCAAGGAGCTTCGCTCTGCAGGAGCTCAAGGAGCTTCGCTCTGCAGGAGGA	830
Qy	864	CRTGAAACCCATGATGCGTGCACCTGCGTGCACCTGCGTGCACAT	917
Db	831	CGTGTACCGTCAAGCTCGCAGGTTGGCCAGGGCTCAACAT	884
RESULT 13			
	AAT29159	standard; cDNA to mRNA; 1055 BP.	
ID	AAT29159	standard; cDNA to mRNA; 1055 BP.	
XX			
AC	AAT29159;		
XX			
DT	16-OCT-2003	(revised)	
DT	18-NOV-1996	(first entry)	
XX			
DE	Carbonyl reductase gene.		
XX			
KW	carbonyl reductase; R-gamma-substituted-beta-hydroxybutyrate;		
KW	gamma-substituted acetocetate; raw material; drug synthesis;		
KW	agricultural chemical production; ds.		
XX			
OS	Sporidiobolus salmonicolor; IP01038.		
XX			
Key	Location/Qualifiers		
CDS	30..1001		
FT	/tag= a		
FT	/product= "carbonyl_reductase"		
FT	33..999		
FT	/tag= b		
XX			
PN	JP08103269-A.		
XX			
PD	23-APR-1996.		
XX			
PF	07-OCT-1994;	94JP-00244030.	
XX			
PR	07-OCT-1994;	94JP-00244030.	
XX			
PA	(BLED) DENKI KASAKU KOGYO KK.		
XX			
DR	WPI; 1996-253873/26.		
P-PSDB;	AAR6294.		
XX			
PT	Carbonyl reductase gene and protein - used in the production of R-Gamma-		
PT	substituted-beta-hydroxybutyrate; for use in drugs and agrochemicals.		
XX			
Claim 1:	Page 9-10; 14pp; Japanese.		
XX			
CC	The present sequence encodes a carbonyl reductase (ALD) isolated from		
CC	Sporidiobolomyces salmonicolor IP01038. The enzyme is used in the synthesis		
CC	of an (R)-gamma-substituted-beta-hydroxybutyrate (GSBH) from a gamma-		
CC	substituted acetocetate. The GSBH is useful as a raw material for the		
CC	synthesis of drugs and agricultural chemicals. (Updated on 16-Oct-2003 to		
CC	standardise OS field)		
XX			
SQ	Sequence 1055 BP; 224 A; 365 C; 284 G; 182 T; 0 U; 0 Other;		
Query Match	50.4%	Score 199.6; DB 2; Length 1055;	
Best Local Similarity	55.5%	Prod. No. 2.5e+43;	
Matches	463;	Mismatches 0; Mismatches 344;	
Conservative		Indels 27; Gaps 3	
Ov			
	87 GGGGAGACCTATACTGCTGTCAACACTGCCCTGAAGACGGTACCGTSACTTGACTG 146		

AC	AAT29160;	Db	360	GAACCTCTTCCGAAAGGCCAACGACAAGGAGGTAAAGCTGGAGGTCAAGCTCAGCCTCGT 419
XX	16-OCT-2003 (revised)	QY	444	GCCCCATGGCGCTATGGAGAGATTATGGGATCCAGCCGGTCCATTTGGGTT 503
DT	18-NOV-1996 (first entry)	DB	420	CGACAGTGGAAAGG-GATGTCAGCTCTGCAGCTTCGAAAGTCAAGGATCGGGT 479
DE	Carbonyl reductase coding sequence.	QY	504	CTCCACTGGACCATGGCCGACCTGCAAGTCCCAAAGTCAAGTCCCAAAGTCAAGTCC 563
XX	carbonyl reductase; R-gamma-substituted-beta-hydroxybutyrate;	DB	480	TTCGACTTGAGGGTACGGAGATGGTACGGGTCAGGAGCTACTGCCT 539
KW	gamma-substituted acetocetate; raw material; drug synthesis;	QY	564	CGCCAACAGCATGGAGATCACCCCTTCCTGCCAACAGGAGGTGGTCAGTACTGCCT 623
KW	agricultural chemical production; ds.	DB	540	CGTCARCRAGATGCGTACCCCTTCCTCCAGGCGACTCGCCACACAA 599
XX	sporidiobolus salmonicolor; IF01038.	QY	624	CTCCAGAGATTATGCCCTGGCTTACCTCTGGCTCGAGAACAGGTTCCAC 683
FH	Location/Qualifiers	DB	600	GGCCAGAAACATTACATTACCTCCCTCGTAACACAGTGTCCGC 659
KEY	1..972	QY	684	CACCGGTAGGGTCAACCTCTCCCTTCCTCCAGGAA 743
CDS	/*tag= a	DB	660	TCTCTTGTCAAGAACGGTCACTTACGTTCCGAGGATCCCGAAGGGAA 710
FT	/product= "carbonyl_reductase"	QY	744	CACCCCTGCTCAGGTCTTATTCGCTGGGTCTCGCCGGTCACTGCTGG 803
FT	4..969	DB	711	CACGCCGGCTAGGTCTCATGGTCTCATGGTCTGCGGCCACTCGT 770
FT	/*tag= b	QY	804	GAGCTCCAATCCCAGCGATTGAGTCGCACTTAAAGGCAATTAGCTTCG 863
XX	JP0103229-A.	DB	771	GTGGTCACCCCTCCGATGGGAGAACTTAAGCAGGCTCGCTCTG 830
XX	23-APR-1996.	QY	864	CTTGAGGCCATCAATGGGTGCCAGGTCTGTCACACTCGTTCGTCAACAT 917
PD	07-OCT-1994;	DB	831	CGTCATGCCGTCAGCAAGTCGGGAGGTTCCGGCAGGGTCAACAT 884
XX	94JP-00244090.	PS		
PR	07-OCT-1994;	PS		
XX	94JP-00244090.	PS		
PA	(ELED) DENKI KAGAKU KOGYO KK.	PS		
XX	WPI; 1996-253873/26.	PS		
DR	P-PDB; AAR96394.	PS		
PT	Carbonyl reductase gene and protein - used in the production of R-gamma-	PS		
PT	substituted-beta-hydroxybutyrate, for use in drugs and agrochemicals.	PS		
XX	Claim 3 ; Page 10-11; 14pp; Japanese.	PS		
XX	Sequence 972 BP; 203 A; 337 C; 265 G; 167 T; 0 U; 0 Other;	PS		
CC	The present sequence encodes a carbonyl reductase (ALD) isolated from	PS		
CC	Sporidiobolus salmonicolor IF01038. The enzyme is used in the synthesis	PS		
CC	of an (R-gamma -substituted-beta-hydroxybutyrate (GSBH) from a gamma-	PS		
CC	substituted acetoacetate. The GSBH is useful as a raw material for the	PS		
CC	synthesis of drugs and agricultural chemicals. (Updated on 16-OCT-2003 to	PS		
CC	standardise OS field)	PS		
XX	Sequence 972 BP; 203 A; 337 C; 265 G; 167 T; 0 U; 0 Other;	PS		
CC	Query Match Score 199.6; DB 2;	PS		
CC	Best Local Similarity 20.4%; Pred. No. 2.5e-43;	PS		
CC	Mismatches 55.5%; Matches 463; Conservative 0; Mismatches 344;	PS		
CC	Indels 27; Gaps 3;	PS		
XX	Sequence 972 BP; 203 A; 337 C; 265 G; 167 T; 0 U; 0 Other;	PS		
QY	20.4%; Score 199.6; DB 2;	PS		
DB	Best Local Similarity 55.5%; Pred. No. 2.5e-43;	PS		
QY	Mismatches 463; Conservative 0; Mismatches 344;	PS		
QY	Indels 27; Gaps 3;	PS		
QY	147 TGCTCGTGTACTACCTGAAACGGGTGAGGTGCTGGTGTGACTCTGTGAGGA 206	PS		
DB	87 GGCGAGACCTATACTGCTCACCACTGCTGAAAGACGGTACCGTCACTGGACTG 146	PS		
DB	75 GGGCAAGTGGCAGGGCTCAAGTCCCATGAGTGTGATACTGTCACCTCGACCT 134	PS		
QY	147 TGCTCGTGTACTACCTGAAACGGGTGAGGTGCTGGTGTGACTCTGTGAGGA 206	PS		
DB	135 TGCRAAGTGTACTGAAACCTGAGGTGCTGGCCATCA 179	PS		
QY	207 GAACCCCTGGTGAAGCCNGAGGACATCTGCTGCAACCTCCGACCTCCA 266	PS		
DB	180 GGAGGCTGGCTCAAGCAGGACCTCTCATACCTCGAACTCGA 239	PS		
QY	267 CGGTATGAGGAGCTCTGGTCACTGAGCTCCCTGAAAGGTGGCTTGACTA 326	PS		
DB	240 CGGCCGGAGCAGGCTGGCCCTGAGGACTCTGCGCTCGAGTA 299	PS		
QY	327 CGTGATACTTCCCTGTTACGCTGCAAGGAAATGCGCAAGGTGAGCC 386	PS		
DB	300 CCTGGACCTTACCTCATGCGGAAATACTGTCATCTCAAG --GACCTGACGAGAACCCGA 443	PS		
QY	387 CAAGATTGGCCCTGACGGAAATACTGTCATCTCAAG --GACCTGACGAGAACCCGA 443	PS		

XX 12-JUN-2002 . DT 22-OCT-2002 (first entry)
 PD XX PCR amplicon of a protein producing (S)-4-bromo-3-hydroxybutanoate.
 PF XX (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;
 PR XX agrochemical; 4-cyano-3-hydroxybutanoic acid; ss.
 PR XX Penicillium citrinum.
 PR XX EP1213354-A2.
 PA XX
 XX (SUMO) SUMITOMO CHEM CO LTD. XX
 PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R; XX
 DR WPI; 2002-550350/59. XX
 PT New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by XX
 PT asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals. XX
 PT and agrochemicals.
 XX PS Example 1; Page 38; 56pp; English.
 XX SQ The present sequence represents a PCR amplified fragment of a gene
 CC encoding a Penicillium citrinum protein. This protein is capable of XX
 CC producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4- XX
 CC bromo-3-oxobutanoate. The amplified fragment was inserted into plasmid XX
 CC pBR-2. The protein and polynucleotides are useful for producing optically active (S)-4-bromo-3-hydroxybutanoate, which is useful as an intermediate CC in the production of pharmaceuticals and agrochemicals. The (S)-4-bromo-3- CC hydroxybutanoate produced can also be used to produce 4-cyano-3- CC hydroxybutanoic acid.
 XX SQ Sequence 743 BP; 183 A; 202 C; 156 G; 167 T; 0 U; 35 Other;
 CC Query Match 42.1%; Score 412; DB 6; Length 743;
 CC Best Local Similarity 100.0%; Pred. No. 1-pe-10;
 CC Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC Db 567 CAACCGATGAGATCACCCCTTCTGGCCAACCGAGACTACTGCTCTC 626
 CC 1 CAACCGATGAGATCACCCCTTCTGGCCAACCGAGCTGTCACTGCTTC 60
 CC Db 627 CAAGAACATATTGCCGTGGCTACTCTCTGGCTGCCAGACAGTTCCACAC 686
 CC 61 CAAGAACATATTGCCGTGGCTACTCTCTGGCTGCCAGACAGTTCCACAC 120
 CC Db 687 CGTAGCGGGTCAGGAGAACATGCGCAAGGGCCAAACAC 746
 CC 121 CGGTGAGGGTCAGGAGAACATGCGCAAGGGCCAAAC 180
 CC Qy 747 CCTGGCTCAAGTTCTTAATGCTGGGTCTGGGTGAGCTGGTCTCCC 806
 CC Db 181 CCTTGCTGAGTTCTATGGTGGTGGTGAAGTCTGGTCTCCC 240
 CC Db 807 CTCCAACCCAAAGGGCATGTCAGTCAGGATGAGCTTCGGATGCGACTT 866
 CC Db 241 CTCCAACCCAAAGGGCATGTCAGTCAAGGATGAGCTTCGGATGCGACTT 300
 CC Qy 867 TGAGGCCTCATGCGCTGCCAAGGCTCGTCAACTTCCCTCAATGAGGATC 926
 CC Db 301 TGAAGGCCATCAATGCCGTGGCTCAAGGGTCTGAGGATGAGGATC 360
 CC Qy 927 TTTCGGATATAATGTTGCCGGAGAGCCCAAGAACCTGCTGCTGA 978
 CC Db 361 TTTCGGATATAATGTTGCCGGAGAGCCCAAGAACCTGCTGCTGA 412
 CC Db 307 TGACCAAGGGTGGACCACTCC 265
 CC Db 307 TGACCAAGGGTGGACCACTCC 331

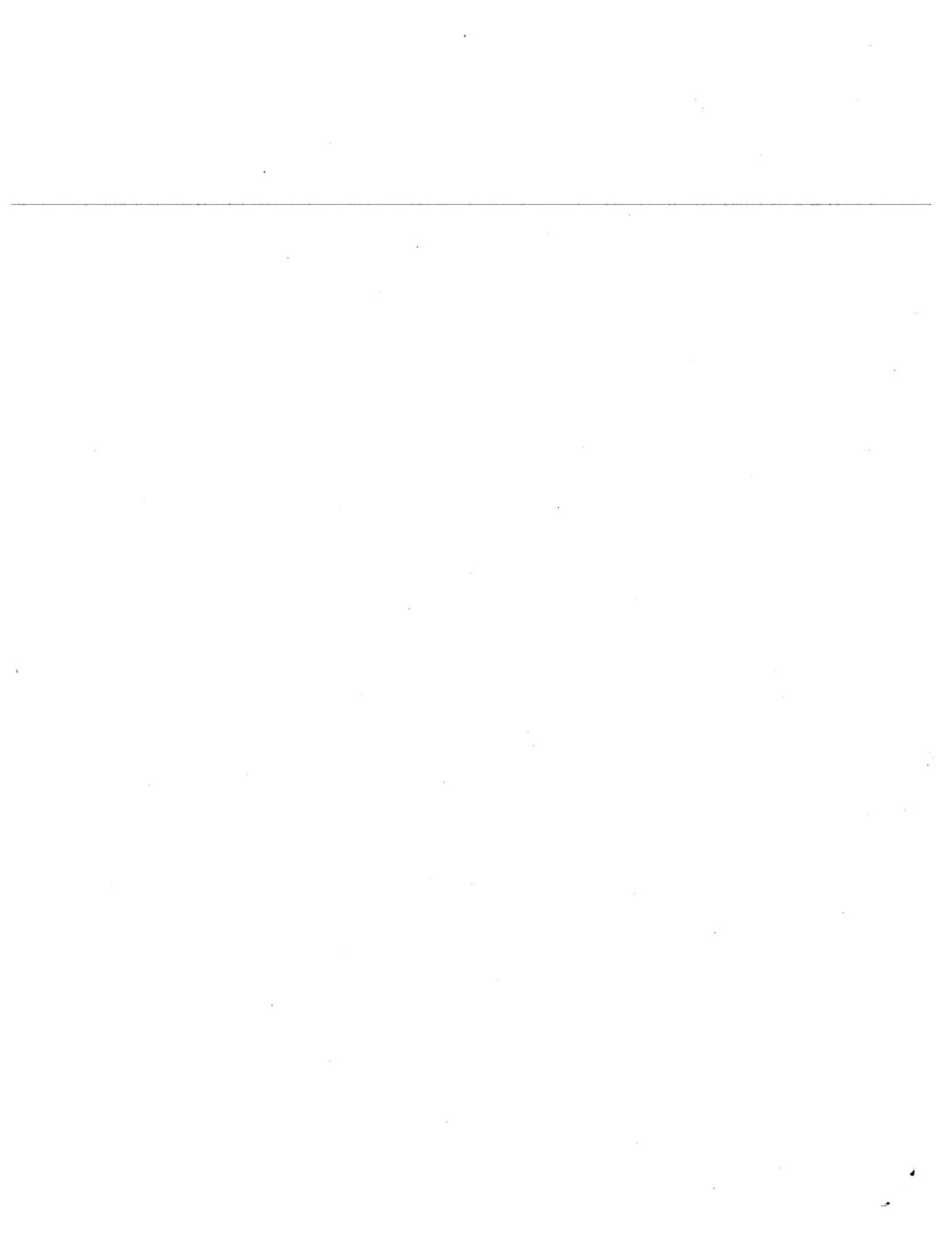
RESULT 11
 AAT29160
 ID ABL5387 standard; DNA; 331 BP.
 XX AC
 XX

RESULT 11
 AAT29160
 ID ABL5387 standard; cDNA to mRNA; 972 BP.
 XX

Query Match	42.6% ; Score 417; DB 6; Length 417;	CC oligotrophic, solid, early germination, alkaline, high temperature, low temperature or maltose culture or polyamino acids stringently hybridising to these sequences. The Polynucleotides are useful for monitoring the progress of fermentation and the growth conditions of a fungus, especially of Aspergillus oryzae which is widely used in industrial fermentation. Also monitoring for fungal contamination. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Best Local Similarity	100.0% ; Pred. No. 6.9e-102;	CC
Matches	417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	CC
Qy	245 CCAAGGTGTTGAAACCCTCACCGTTATGGAGACGTCTGGTCCATTGAGACTCCC 304	CC
Db	1 CCAAGGTGTTGAAACCCTCACCGTTATGGAGACGTCTGGTCCATTGAGACTCCC 60	CC
Qy	305 TGAACCGTCTGGACTACCTGATAATGTCCTGTTCACTGGCCATTGCTGCCG 364	CC
Db	61 TGAAGC GTCTGGACTACCTGATAATGTCCTGTTCACTGGCCATTGCTGCCG 120	CC
Qy	365 AGAGAAATGCCAACGGTGAGCCAAATACTGTCATTCTCAAGG 424	CC
Db	121 AGAGAAATGCCAACGGTGAGCCAAATACTGTCATTCTCAAGG 180	CC
Qy	425 ACCTGACCGAAACCCGAGCCCATGGCCGCTATGGAGAAGATTATGAGATGCCA 484	CC
Db	181 ACCTGACCGAAACCCGAGCCCATGGCCGCTATGGAGAAGATTATGAGATGCCA 240	CC
Qy	485 AGGCCAGGTCATCTGGTCAACTGGACCATGGCCAACTTGAGAGATGTCAGT 544	CC
Db	241 AGGCCAGGTCATCTGGTCAACTGGACCATGGCCAACTTGAGAGATGTCAGT 300	CC
Qy	545 TGCCTAAGGTCAATGCTCAAGCCAACTGAGATTCAACCTCTGGCCAAACGAGG 604	CC
Db	301 TGCCTAAGGTCAATGCTCAAGCCAACTGAGATTCAACCTCTGGCCAAACGAGG 360	CC
Qy	605 AGCTGTGCACTGTGCTCTCCAGAAACATTATGCCCTGGCTACTCTCTCTGG 661	CC
Db	361 AGCTGTGCACTGTGCTCTCCAGAAACATTATGCCCTGGCTACTCTCTGG 417	CC
RESULT 8		
ID ABZ53802	standard; cDNA; 587 BP.	
XX ABZ53802/		
AC		
XX DT 28-MAR-2003 (first entry)		
XX Aspergillus oryzae polynucleotide SEQ ID NO 2915.		
DE		
XX Aspergillus oryzae: fermentation; fungus; industrial; EST;		
KW expressed sequence tag; gene; ss.		
XX Aspergillus oryzae.		
OS WO200279476-A1.		
XX PD 10-OCT-2002.		
XX PP 22-MAR-2002; 2002WO-IB0000890.		
XX PR 30-MAR-2001; 2001JP-00098371.		
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.		
PA (NAEB-) NAT RES INST BREWING.		
PA (NORQ) NAT FOOD RES INST MIN AGRIC.		
XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H, Abe K;		
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K,		
XX DR 2003-046817/04.		
XX PT Detection of expression of specific Aspergillus genes for monitoring the fermentation and growth conditions of the fungus, using DNA probes.		
PT Claim 1; SEQ ID NO 2915; 48pp + Sequence Listing; Japanese.		
XX CC The invention relates to a polynucleotide having any of 606 specific sequences (ABZ50880 ABZ56893), which are expressed by a fungus under specific culture conditions including one or more of eutrophic,		
CC CC		
XX XX	RESULT 9	
XX ID ABL59386		
XX TD ABL59386 standard; DNA; 743 BP.		
XX AC ABL59388;		
XX DT 22-OCT-2002 (first entry)		
XX DE PCR amplicon of a protein producing (S)-4-bromo-3-hydroxybutanoate.		
XX KW (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical;		
KW agrochemical; 4-cyano-3-hydroxybutanoic acid; ss.		
XX OS Penicillium citrinum.		
XX PN EP1213354-A2.		

Db	301	TCCCTGAAAGCCTTGGACTTGACATGCTGATACTGGTTAACGGCCATTGGT	360	
Qy	361	GCGGAGAAATGGCCGGTGAACCAAGATTGGCTTACGGCTATTCTGC	420	PR 07-DEC-2000; 2000JP-00372704.
Db	361	GCGGAGAAATGGCCGGTGAACCAAGATTGGCTTACGGCTATTCTGC	420	PR 15-JAN-2001; 2001JP-000614.
Qy	421	AAGGACCTGACGGAGAACCCCAGGCCCATGGCGCTATGGAAAGTTTGGAGAT	480	PR 02-FEB-2001; 2001JP-00026514.
Db	421	AAGGACCTGACGGAGAACCCCAGGCCCATGGCGCTATGGAAAGTTTGGAGAT	480	PR 11-JUN-2001; 2001JP-00175175.
Qy	481	CCCAAGGCCAGGTCCATTGGTCTCCAATGGACCATTCGCAACTTGGTCC	540	XX (SUMO) SUMITOMO CHEM CO LTD.
Db	481	CCCAAGGCCAGGTCCATTGGTCTCCAATGGACCATTCGCAACTTGGTCC	540	PA
Qy	541	Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;	XX	PI
Db	541	Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;	XX	XX
Qy	541	WPI ; 2002-5501350159.	DR	DR
Db	541	DR-P-PSD; ABB77965.	DR	DR
Qy	541	New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals and agrochemicals.	XX	XX
Qy	541	Example 1; Page 43-45; 56pp; English.	PS	PS
Db	541	AAGTTCSCCAAGCTCATCCCTACGGCAACAGATCAGATTCAACCCCTTGGCCAC	600	XX
Db	541	AAGTTCSCCAAGCTCATGGCTTCAGCAACAGATCAGATTCAACCCCTTGGCCAC	600	CC
Qy	601	GAGGAGTTGGTCACTGGTCAAGACATATGGCGGACCTTCGACTCCTCTCG	660	CC
Db	601	GAGGAGTTGGTCACTGGTCAAGACATATGGCGGACCTTCGACTCCTCTCG	660	CC
Qy	661	GGCTCGGAACCAAGGTTCCACCAACGGTGAAGGACTGTGAC	720	CC
Db	661	GGCTCGGAACCAAGGTTCCACCAACGGTGAAGGACTGTGAC	720	CC
Qy	721	GAGATCSCGAAGGGGAACACCCCTGGCTTATTGGCTGGGTCTGCC	780	CC
Db	721	GAGATCCCGAAAGGGGAAACCCCTGGCTTATTGGCTGGGTCTGCC	780	CC
Qy	781	CTGGCTTAACGGCTTCTCCAAAGAGCTCCAAACCTTCAG	840	CC
Db	781	CTGGCTTAACGGCTTCTCCAAAGAGCTCCAAACCTTCAG	840	CC
Qy	841	AGCATTTGAGCTCTCGATGGGACTTTGANGCCATCAATGGCGTTGCCAAGGGTGTCTAC	900	CC
Db	841	AGCATTTGAGCTCTCGATGGGACTTTGANGCCATCAATGGCGTTGCCAAGGGTGTCTAC	900	CC
Qy	901	TTCGGTTCTGGCAACTGAAAGATAATTGGATGATGTTCTGCCAGGAGACGCC	960	CC
Db	901	TTCGGTTCTGGCAACTGAAAGATAATTGGATGATGTTCTGCCAGGAGACGCC	960	CC
Qy	961	AAGAACCTGTCGGTCAACATGAAAGATACTTCTGGATATGATGTTCTGCCAGGAGACGCC	978	CC
Db	961	AAGAACCTGTCGGTCAACATGAAAGATACTTCTGGATATGATGTTCTGCCAGGAGACGCC	978	CC
RESULT 4				
ID	ABL59397	standard; DNA; 996 BP.		
ID	ABL59397	standard; DNA; 996 BP.		
XX				
AC	ABL59397;			
XX				
AC	ABL59397;			
XX				
DT	22-OCT-2002 (first entry)			
XX				
DE	DNA sequence of a protein producing (S)-4-bromo-3-hydroxybutanoate.			
XX				
KW	(S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; pharmaceutical; agrochemical; 4-cyano-3-hydroxybutanoic acid; gene; ss.			
XX				
OS	Penicillium citrinum.			
XX				
PH	Location/Qualifiers			
CDS	1..978			
FT	*tag= a			
XX				
EP1213354-A2.				
XX				
PD	12-JUN-2002.			
XX				
PP	07-DEC-2001; 2001EP-00310251.			
XX				
Qy	601	GAGGAGCTGGTGGAGTACTGGCTTCTCAAGAACATTATGCCGTGCTACTCTCTCTG	660	

XX OS Penicillium citrinum.
 XX Key Location/Qualifiers
 PH 1. *tag= a
 CDS /product= "Penicillium reducing enzyme protein"
 FT XX
 FT FT
 FT XX
 XX BP1321827-A2.
 XX PD 02-JUL-2003.
 XX 20-DEC-2002; 2002EP-00258814.
 XX PR 27-DEC-2001; 2001JP-00395884.
 XX PR 27-DEC-2001; 2001JP-00395885.
 XX PR 10-APR-2002; 2002JP-00107648.
 XX (SUMO) SUMITOMO CHEM CO LTD.
 PA Asako H, Wakita R, Itoh N;
 PA WPI: 2003-723302/69.
 DR P-PSDB; ADE39629.
 XX Producing optically active 2-hydroxycycloalkanecarboxylic acid ester by
 PR reacting 2-oxocycloalkanecarboxylic acid ester with transformant having
 PR ability to assymmetrically reduce the ester to optically active ester.
 XX Claim 9; SEQ ID NO 4; 45pp; English.
 XX The invention relates to a novel method for producing an optically active
 CC 2-hydroxycycloalkanecarboxylic acid ester. The novel method involves:
 CC allowing 2-oxocycloalkanecarboxylic acid ester to react with a
 CC transformant, or a dead cell or their extract, artificially provided with
 CC the ability to assymmetrically reduce a 2-oxocycloalkanecarboxylic acid
 CC ester to the optically active 2-hydroxycycloalkanecarboxylic acid; and an
 CC ability to regenerate a coenzyme on which an enzyme having the above
 CC ability resides, and collecting the resulting optically active 2-
 CC hydroxycycloalkanecarboxylic acid. The optically active 2-
 CC hydroxycycloalkanecarboxylic acid is useful as an intermediate for the
 CC production of bioactive substances. This polynucleotide sequence
 CC represents the DNA encoding a Penicillium enzyme used to reduce 2-
 CC -oxocycloalkanecarboxylic acid ester to the optically active 2-
 CC hydroxycycloalkanecarboxylic acid as part of the method of the invention.
 XX Sequence 978 BP; 218 A; 282 C; 261 G; 217 T; 0 U; 0 Other;
 Query Match 100.0%; Score 978; DB 9; Length 978;
 Best Local Similarity 100.0%; Pred. No. 1.7e-252; Indels 0; Gaps 0;
 Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCTAACGGAAAGACTTCACTTGGACTTTCAATGAGCTTCAAGATTCTGGCTTT 60
 Db 1 ATGCTAACGGAAAGACTTCACTTGGACTTTCAATGAGCTTCAAGATTCTGGCTTT 60
 QY 61 CGATGGTACCGGTTCTCCAAAGAGCTCAACCCAAAGGCTTCAAGCTTCAACTTCAG 840
 Db 781 CGTGGTACCGGTTCTCCAAAGAGCTCAACCCAAAGGCTTCAAGCTTCAACTTCAG 840
 QY 841 AGCATTTGAGCTCTCCGATGGGACTTTGAGCCATTAACTCGGTGCTTCAAGGGTGTCAC 900
 Db 841 AGCATTTGAGCTCTCCGATGGGACTTTGAGCCATTAACTCGGTGCTTCAAGGGTGTCAC 900
 QY 901 TTCCGTTTCGTCACATGAAGGATACTTGGATGATGATGATGTCGGCCAGGACCCGC 960
 Db 901 TTCCGTTTCGTCACATGAAGGATACTTGGATGATGATGTCGGCCAGGACCCGC 960
 QY 961 AAGAAGCTGCTGCCNGA 978
 Db 961 AAGAAGCTGCTGCCNGA 978
 RESULT 3
 ADE39630 standard; DNA; 978 BP.
 XX QY 181 GGGGGATTCGTAACCTCGTGAAGGGAAACCCCTCGTGAAGGGCATCTTCGTC 240
 AC XX
 DT 29-JAN-2004 (first entry)
 XX Db 181 GGGGGATTCGTAACCTCGTGAAGGGAAACCCCTCGTGAAGGGCATCTTCGTC 240
 DE Penicillium reducing enzyme protein encoding DNA, SEQ ID No 4.
 XX QY 241 TGACCAAGGTGGACCCACCTCCACCGTTATGGGAGTCGTCCTGGTCAATGAGCAG 300
 KW 2-hydroxycycloalkanecarboxylic acid ester; enzyme; Penicillium; reduction;
 gene; ds
 QY 301 TCCCTGAAAGCGCTCTGGACTGACTAGTTGATATGTTCCCTCGTGTACITGCCATTTGCT 360



Qy	362	CCGAGAAGAAATGGCCAGGTTGCCCCAAGATTGGCCCTTGACGCCRAATAACGTCAATTCTCA	421
Db	245	CGGAGAGAACGAGACTATGAGTCAAGATCGAGATGATGAAACTACATCATACA	304
Qy	422	AGGACCTGACCGAGAACCCGAGCCACATGGGGCTATGGAAAGATTTATGGAGATC	481
Db	305	AAGAGTCAGGCCAACATGGACCTATGGGTAGATTCAGGGCTTCAAAAGCTG	364
Qy	482	GCAAGGCCAGGTCCATTGGTCTCCAACTGGGACATTGCCAACCTTGAGAGATGTC	541
Db	365	GCAAGGCCAACGGCATCGAAATCCAAATTGACATCTCAAACCTGAAAGTC	424
Qy	542	AGTTGCCAAGGTGATGCCAACCCAACTGAGATCACCCCTTGCCAAAG	601
Db	425	AGTAGCAGCATGTTCTCCGCCATCACCAAGTCGAGATCATCCGGTTCGGCTAAC	484
Qy	602	AGGACCTGGTGTGAGTACTCTCTCCAAAGAACATTATCCCGGGCTACTCTCTGG	661
Db	485	CTAAGTTGATCAAATCTACTCTTTCAANGAACATCCTTCCGGCATATGGCACTGG	544
Qy	662	GCTGCCAGAACCAAGTTCCACACCGGTTGAGGGTAAGCGAGAACAGACTCTGAACG	721
Db	545	GATCCCAGAGCCAGTGCCCAAGCAGGAAACCGTATCCTGAAACTCTGAACTPATCT	604
Qy	722	AGATGCCAGAACGGCCAAACACCCCTGCTAGGTTCTATGGCTGGGTCTGCC	781
Db	605	TAATTCGGAAAAGAGGGCGTTAGCATAGTCAAATTGATGCTTGTTATCAAC	664
Qy	782	G-TGGCTACCTCGTTCTCCCAAGGTCACCCCAAGGCTGAGTGGATGCTCAACTCAAG	840
Db	665	GAAGGATACCTAACCTTCCATGAGTCAAATGGGGGAGATCAAAACCAATGGAAACA	724
Qy	841	AGCAATGAGGCTCGGATCGAAGGCAATCATGGCTTGCCTTGCCTAAG	891
Db	725	CTTGTGATTGACTGCAAGGGTTGAAAATGACCCAAAGTCCGAGG	775

FEATURES	Higashihiroshima, Hiroshima 739-0046, Japan (E-mail:akao.t@nrib.go.jp, Tel:81-824-20_0824, Fax:81-824-20_0808)	Qy	721 GAGATGCCGAGAACCGGGAAACCCCTGCCTGAGTTTATTGCCCTGGGGTCGCC 780
source	Location/Qualifiers 1..1183 /organism="Aspergillus oryzae" /mol_type="mRNA" /strain="RIBA0" /db_xref="txxon:5062" /clone="JK1125" 53 ..1040 /codon_start=1 /product="reductase-like protein" /protein_id="BAC56099.1" /db_xref="GI:28144501" /translation="MSNGKTFITLNSNGKIPGPGFTASESGSKGETYQAVTKALETY RHDCAYNLYNEEVGOGIDELKKNSVREDFIYCTKVNHLRPEDVWSIDNSL GRKRLDYLDFLWHMPIASERKDQERPKIPGKVKYLKELETPPLNNELVCPPTWAMEKLYRD PLOSSQNPVPTGERUSENKTNLNAQKGCHTLAQVLIANGRRGTVLPSSNPAPRIE SNFKSIELTDEDEFAVNRAEGRHFRFVNMDHFRGTDVWPBETAKNLSA"	Db	773 GAGATGCCGAGAACCGGGAAACCCCTGCCTGAGTTTATTGCCCTGGGGTCGCC 832
CDS		Qy	781 CGGGCTACCTGCTCTCCCAAGAGCTCAAGCTCAACTTCAACTCTGCTGCT 840
		Db	833 CGGGCTACGGTGTGTCCTCCAAAGGCTCAACCCAAAGCTCAAGCTCAACTCTGCTGCT 892
		Qy	841 AGCATGGACTCTCCATGCCACTTGAAGCCATCAATGCCCTGCCAAGGGTCGTAC 900
		Db	893 AGCATGGACTCTCCATGCCACTTGAAGCCATCAATGCCCTGCCAAGGGTCGTAC 952
		Qy	901 TICCGTTTCGTCACATGAGGATACTTGAGATATGATGCTGCCAGGAGACGCC 960
		Db	953 TICCGTTTCGTCACATGAGGATACTTGAGATATGATGCTGCCAGGAGACGCC 1012
		Qy	961 AAGAACCTGCTGCTGCTG 978
		Db	1013 AAGAACCTGCTGCTGCTG 1030
ORIGIN			
		Query Match Score 734.8; DB 8; Length 1183;	RESULT 6
		Best Local Similarity 84.5%; Pred. No. 1..9e-147; Indels 0; Gaps 0;	ANI496025
		Matches 826; Conservative 0; Mismatches 152;	LOCUS
			Aspergillus nidulans gldB gene for NADP(+)-dependent glycerol
			DEFINITION dehydrogenase, exons 1-2.
			ACCESSION AJ496025
			VERSION AJ496025.1 GI:32879319
			KEYWORDS gldB gene; NADP (+)-dependent glycerol dehydrogenase;
			ORGANISM Aspergillus nidulans (anamorph: Aspergillus nidulans)
			Emericella nidulans (anamorph: Eurotiomycetes;
			Bakteria; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
			Eurotiales; Trichocomaceae; Emericellaceae.
			1
		REFERENCE de Vries R.P.	REFERENCE
		AUTHORS de Vries R.P.	AUTHORS
		TITLE Unpublished	TITLE
		JOURNAL (bases 1 to 2397)	JOURNAL
		AUTHORS de Vries, R.P.	AUTHORS
		TITLE Direct Submission	TITLE
		JOURNAL Submitted (16-JUL-2002) de Vries R.P., Microbiology, Utrecht University, Padualaan 8, Utrecht, 3584 CH, NETHERLANDS	JOURNAL
		FEATURES	FEATURES
		source /organism="Emericella nidulans"	source
		/product="NADP(+)-dependent glycerol dehydrogenase in Aspergillus nidulans	/product="NADP(+)-dependent glycerol dehydrogenase in Aspergillus nidulans"
		/db_xref="taxon:162425"	/db_xref="taxon:162425"
		gene /gene="gldB"	gene
		/protein_id="CAD42649.1"	/protein_id="CAD42649.1"
		/db_xref="GI:1100,1159,1728"	/db_xref="GI:1100,1159,1728"
		/gene="gldB"	/gene="gldB"
		/EC number="1.1.1.72"	/EC number="1.1.1.72"
		/function="glycerol biosynthesis"	/function="glycerol biosynthesis"
		/codon_start=1	/codon_start=1
		/product="NADP(+)-dependent glycerol dehydrogenase"	/product="NADP(+)-dependent glycerol dehydrogenase"
		/db_xref="MSSGKTFIKLSQEVGDFAVRFLKENPSKREDFIICHTKWNHLHRPDVRSIEDSL RHLDCAWFQNEDEVGDFAVRFLKENPSKREDFIICHTKWNHLHRPDVRSIEDSL KKLKDNYDVLFLHWPIAEKESODDKPKIGPDKPWTQAMKLYED KLARSIGSWNTIEGLIKYAKPHVQIHHPELPNPELJOVQLAWGLRGYVWLPSNSNPARIE PLGSQNQVPTGEVSSENTLINEAQKQGNTLQAVLAWGLRGYVWLPSNSNPARIE SNFCTIUSDEDYAVNAVAKGRHFRTVMKDFFGVDWPEETAKNLSA"	/db_xref="MSSGKTFIKLSQEVGDFAVRFLKENPSKREDFIICHTKWNHLHRPDVRSIEDSL RHLDCAWFQNEDEVGDFAVRFLKENPSKREDFIICHTKWNHLHRPDVRSIEDSL KKLKDNYDVLFLHWPIAEKESODDKPKIGPDKPWTQAMKLYED KLARSIGSWNTIEGLIKYAKPHVQIHHPELPNPELJOVQLAWGLRGYVWLPSNSNPARIE PLGSQNQVPTGEVSSENTLINEAQKQGNTLQAVLAWGLRGYVWLPSNSNPARIE SNFCTIUSDEDYAVNAVAKGRHFRTVMKDFFGVDWPEETAKNLSA"
		exon 693 ..1100	exon 693 ..1100
		/gene="gldB"	/gene="gldB"
		/number=1	/number=1
		intron 1101 ..1158	intron 1101 ..1158
		/gene="gldB"	/gene="gldB"
		/number=1	/number=1
		exon 1159 ..11728	exon 1159 ..11728
		/gene="gldB"	/gene="gldB"

Qy	781	CCTGGCTTACCGGCTTCCTCCAAAGACTCTCCAAAGGCCATGGTCAACTTCAG	840	
Db	781	CGTGGCTTACCGGCTTCCTCCAAAGACTCTCCAAAGGCCATGGTCAACTTCAG	840	
Qy	841	AGCATTTGAGCTTCCATGGGACTCTTGAACCATTAATGGCTTCAAGGTCTCAC	900	
Db	841	AGCATTTGAGCTTCCATGGGACTCTTGAACCATTAATGGCTTCAAGGTCTCAC	900	
Qy	901	TTCGGTTTCGTCACATGAAGGATACTTTCGGATATGATGTCCTGGCCAGGGAGCC	960	
Db	901	TTCGGTTTCGTCACATGAAGGATACTTTCGGATATGATGTCCTGGCCAGGGAGCC	960	
Qy	961	AAGAACCTGTCGTCACATGAAGGATACTTTCGGATATGATGTCCTGGCCAGGGAGCC	960	
Db	961	AAGAACCTGTCGTCACATGAAGGATACTTTCGGATATGATGTCCTGGCCAGGGAGCC	960	
RESULT 4				
	AX472813	AY472813	996 bp	DNA
DEFINITION	Sequence 28	from Patent	EP1213354.	
ACCESION	AY472813			PAT 09-AUG-2002
VERSION	AY472813.1	GI:22207663		
KEYWORDS				
SOURCE	Penicillium citrinum			
ORGANISM	Penicillium citrinum			
Eukaryot; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Penicillaceae; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.				
REFERENCE				
AUTHORS	Asako,H., Matsumura,K., Shimizu,M., Ito,N. and Wakita,R.			
TITLE	Process for producing optically active 4-halo-3-hydroxybutanoate			
JOURNAL	Patent: EP 1213354-A 28 12-JUN-2002;			
FEATURES	Sumitomo Chemical Company, Limited (JP)			
source	Location/Qualifiers			
	-1..996			
	/organism="Penicillium citrinum"			
	/mol_type="unassigned DNA"			
CDS	/db_xref="taxon:5077"			
	1..978			
	/note="unnamed protein product"			
	/protein_id="CAD43582_1"			
	/db_xref="GI:122207864"			
	/db_xref="RENTREMBI:CAD43582"			
	/translation="MSNCKTIFTSAGGAAACTTCAATGGACGAGCTTCAGATTCTGGCTCGCTT"			
	RHDCAWWYLNSEGEGDDEKKEPSVKEELEFCTKWNLNEEDVNLNIDDSL			
	KRGLDDYDMFLWPIAKENGGGEPKGDGXVILKDLTNPMPWRMKEIYD			
	RKRSIGSNWTADELXNSKPKVAMPANOELHPFLPNEEVNOYCFSKNMIVAYS			
	PISQONQVPTPGEPSVENKNTIAEKGNTLQLWLNGRLQYVILPKSSNPKRIE			
	SNFKSIELSDADFEAINAVAKGRHRFVNMDKDFGYDVWEETAKNLSA"			
ORIGIN				
Query Match	100.0%	Score 978;	DB 6;	Length 996;
Best Local Similarity	100.0%	Pred. No. 1..1e-199;		
Matches	978;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
Qy	1	ATGTCCTAACGAAAGACTTCAATGGACGAGCTTCAGATTCTGGCTCGCTT	60	
Db	1	ATGTCCTAACGAAAGACTTCAATGGACGAGCTTCAGATTCTGGCTCGCTT	60	
Qy	61	GGTACCTTCGCTTAGTGAAGGTTCCAAGGGCAGACCTATCTGGCTCACACTGCCCTG	120	
Db	61	GGTACCTTCGCTAGTGAAGGTTCCAAGGGCAGACCTATCTGGCTCACACTGCCCTG	120	
Qy	121	AAGACCGGTACCGTACACTGGACTCTGGTCAAGTCTGGTCAAGGGTGAGSTTGT	180	
Db	121	AAGACCGGTACCGTACACTGGACTCTGGTCAAGGGTGAGSTTGT	180	
Qy	181	GAGGGTATCCTGACTTCAGGAAACCCCTGGTGAAGCTTCCTGGTCAAGGGTGAGSTTGT	240	
Db	181	GAGGGTATCCTGACTTCAGGAAACCCCTGGTGAAGCTTCCTGGTCAAGGGTGAGSTTGT	240	
Qy	241	TGCACCAAAGGTGGAAACCACTCCACCGTTATGAGGACCTCCCTGGTCAATTGAGC	300	
Db	241	TGCACCAAAGGTGGAAACCACTCCACCGTTATGAGGACCTCCCTGGTCAATTGAGC	300	
Qy	301	TCCCTGAGCTCTTGGACTCTGCTGATATGCTTCTCTGTTACTGGCCCATGTCT	360	
Db	301	TCCCTGAGCTCTTGGACTCTGCTGATATGCTTCTCTGTTACTGGCCCATGTCT	360	
Qy	361	GCCGAAAGATGGCAAGGGACTTGAAGGAAATAAGTGTCTTC	420	
Db	361	GCCGAAAGATGGCAAGGGACTTGAAGGAAATAAGTGTCTTC	420	
Qy	421	AAGGACCTGACGGAAACCCGAGCCACATGGGGCTATGGAGAGATTATGGAGGAT	480	
Db	421	AAGGACCTGACGGAAACCCGAGCCACATGGGGCTATGGAGAGATTATGGAGGAT	480	
Qy	481	CGCAAGCCAGGTCCATGGTCAACTGGGACTTGAAGAGATTGTCCTCC	540	
Db	481	CGCAAGCCAGGTCCATGGTCAACTGGGACTTGAAGAGATTGTCCTCC	540	
Qy	541	AAGTGGCCAAAGGTCTGGCTCACGCCAACAGATGGAGATTCAACCCCTCTGCCAAC	600	
Db	541	AAGTGGCCAAAGGTCTGGCTCACGCCAACAGATGGAGATTCAACCCCTCTGCCAAC	600	
Qy	601	GAGGACTCTGGTCAAGTACTGTTCTCAAGAACATATGCCGTGCCCTAATCTCTCTG	660	
Db	601	GAGGACTCTGGTCAAGTACTGTTCTCAAGAACATATGCCGTGCCCTAATCTCTCTG	660	
Qy	661	GGCTCCAGAACGAGCTCCACCCACGGTCCACCGGTCAGGGTCAACCTGAAAC	720	
Db	661	GGCTCCAGAACGAGCTCCACCCACGGTCCACCGGTCAGGGTCAACCTGAAAC	720	
Qy	721	GAGATGCCGAGAACCCCTCCACCCACGGTCCACCGGTCAGGGTCAACCTGAAAC	780	
Db	721	GAGATGCCGAGAACCCCTCCACCCACGGTCCACCGGTCAGGGTCAACCTGAAAC	780	
Qy	781	CGTGGCTACGTCGCTTCCTCCCAAGACGCTCAACCCAAAGCTCCACACTCAAG	840	
Db	781	CGTGGCTACGTCGCTTCCTCCCAAGACGCTCAACCCAAAGCTCCACACTCAAG	840	
Qy	841	AGCATTTGAGCTCTCCATGGCACTTGAGCCATTAATGCCGTGCAAGCTTCAC	900	
Db	841	AGCATTTGAGCTCTCCATGGCACTTGAGCCATTAATGCCGTGCAAGCTTCAC	900	
Qy	901	TTCCGGTTTCGTCACATGGACGATTCCTGGATATGATGTCCTGCCGAGGAGACGCC	960	
Db	901	TTCCGGTTTCGTCACATGGACGATTCCTGGATATGATGTCCTGCCGAGGAGACGCC	960	
Qy	961	AAGAACCTGTCGCTGCGTGA	978	
Db	961	AAGAACCTGTCGCTGCGTGA	978	
RESULT 5				
AB079030		1183 bp	mRNA	lindax
LOCUS				PIN 30-JAN-2003
DEFINITION				Aspergillus oryzae mRNA for reductase-like protein, complete cds.
ACCESSION				
VERSION				AB079030.1 GI:28144500
KEYWORDS				
SOURCE				
ORGANISM				
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Aspergillus oryzae				
EUROTIALES; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus oryzae				
unpublished				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				

FEATURES	Sumitomo Chemical Company, Limited (JP)
source	1 .. 978 /organism="Penicillium citrinum" /mol_type="unassigned DNA" /db_xref="taxon:5077" 1 .. 978 /note="unnamed protein product" /codon_start=1 /protein_id="CAD41580_1" /db_xref="GI:2207641" /db_xref="RENTREMBL:CAD43580" /translation="MSNQKTFLSNGVTKPGVFGTFASEGGKGETYAVTALKGY RHLDCAWVYLINEGEVGIRDEPDKENSKVREDIFUVTKVNHLRVDILWSDSL KRUGLDYVDMFLYWPMLAERKQGPKVILKDUDTENPBTWAKYED RKSRSIGNQPTTADLERNSKPAKXMPANQEITHPLPNELVOYCFSKNIMPAYS PLGSQSONQPTTGRVSЕНКTLNEAЕGGNTLAQULJAWLGRGVYLPSNSPKIE SNFSKIELSDADFBAINAVAKGRHFRFVNMDIFGDWPEETAKNLSA"
CDS	
ORIGIN	
Query Match	100.0% ; Score 978; DB 6; Length 978; Best Local Similarity 100.0%; Pred. No. 1e-19; Gaps 0; Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGCTAACGAAAGACTTACATTGAGAACCGGTCAAAGATTCCCTGGTGGCTT 60
Db	1 ATGCTAACGAAAGACTTACATTGAGAACCGGTCAAAGATTCCCTGGTGGCTT 60
Qy	61 GGTACCTTCGTTAGTGAAAGGTCCAAGGGGAGACCTATACTGGTGTGCCCTG 120
Db	61 GGTACCTTCGTTAGTGAAAGGTCCAAGGGGAGACCTATACTGGTGTGCCCTG 120
Qy	121 AAGACCGGTACCTGTTACCTGACTCTGGTGTGCCCTGACTACTGAAAGGGTGGCTT 180
Db	121 AAGACCGGTACCTGACTCTGGTGTGCCCTGACTACTGAAAGGGTGGCTT 180
Qy	181 GAGGGTATCGTGGACTCTCTGGTGAAGGTGAGGACATCTTCCTC 240
Db	181 GAGGGTATCGTGGACTCTCTGGTGAAGGAACCCCTGGTGAAGGTGAGGCTT 240
Qy	241 TGCACCAAGGTGTGAAACCACTCTCCACCGTTATGGACCTCTGGGCCATGGAC 300
Db	241 TGCACCAAGGTGTGAAACCACTCTCCACCGTTATGGACCTCTGGGCCATGGAC 300
Qy	301 TCCCTGAAAGCCTCTGGTGAATCTGGTGTGATGTTCTCGTCATGGCCATTCCT 360
Db	301 TCCCTGAAAGCCTCTGGTGAATCTGGTGTGATGTTCTCGTCATGGCCATTCCT 360
Qy	361 GCCGAAAGAATGGCAAGGTGAGCCAAAGATGCCCTTAAGGGAAATAATAGTCAATTCTC 420
Db	361 GCCGAAAGAATGGCAAGGTGAGCCAAAGATGCCCTTAAGGGAAATAATAGTCAATTCTC 420
Qy	421 AAGGACCTGACCGAACCCGAGGCCAACATGGCCGTTAGGAAAGTTATGAGGT 480
Db	421 AAGGACCTGACCGAACCCGAGGCCAACATGGCCGTTAGGAAAGTTATGAGGT 480
Qy	481 CGCAAAGGCAGGTCCATTGGGCAATGGGACATTCGCGACCTTGAGAAGATGTC 540
Db	481 CGCAAAGGCAGGTCCATTGGGCAATGGGACATTCGCGACCTTGAGAAGATGTC 540
Qy	541 AAGTTCGCAAGGTGTGGCTTACGCCAACAGATGGGATTCCCTGCCAAC 600
Db	541 AAGTTCGCAAGGTGTGGCTTACGCCAACAGATGGGATTCCCTGCCAAC 600
Qy	601 GAGGACTGGTGGCAGTACTGGTCTCCACCAATATGCCGCTTACCTCTGGCTCTCTG 660
Db	601 GAGGACTGGTGGCAGTACTGGTCTCCACCAATATGCCGCTTACCTCTGGCTCTCTG 660
Qy	661 GGCTCGAGAACAGTGTCCACCACTGGTGTGGGAGAACAGACCTGAC 720
Db	661 GGCTCGAGAACAGTGTCCACCACTGGTGTGGGAGAACAGACCTGAC 720
Qy	721 GAGATCGCGAGAACGCCAGGTCCACCCACACCGGTGAGCTGAGAACCCCTCGTGAAGGGAAACCCCTCGTGAAGGGAAACCCCTCGTGAAGGGAAATCTCGTC 780

Db 230 :|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
EDPKIKSTAAKHKKTAQVLTRFHICRNVITIPKSVTPARLVENTQVFDEKLSDDEEMATI 289
Qy 295 VAKGRHFRFVNMKDTEFYDVWP 316
Db 290 LSFNRNWRACHNVLQSSHLEDYP 311

Search completed: May 25, 2004, 03:45:09
Job time : 48 secs

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 / TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 / FILE REFERENCE: 21:0121_478C14
 / CURRENT APPLICATION NUMBER: US/09/702,705
 / NUMBER OF SEQ ID NOS: 1833
 / SOFTWARE: FastSEQ for Windows Version 3.0
 / SEQ ID NO: 340
 / LENGTH: 316
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-702-705-340

Query Match 28.1%; Score 487; DB 4; Length 316;
 Best Local Similarity 36.6%; Pred. No. 8.8e-44;
 Matches 118; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

Qy 6 TFT-LSGVVKIPGVGFTFASEGSKGEVTYAVTALKTGRLDCAWYVLINEGEVGIR 64
 Db 3 TFEVLSTKAKNPPIVGLGTWKS - PLGKVKEAVKVAIDAGTRHIDCAYVYNEHEYGEAQ 60

Qy 65 DFLKENPSVKREDIFVCTKVNHLHYEDVLSIDSLSLRGLDVTMFLVHMPAAEKN 124
 Db 61 EKIQEK-AVKREDLFTVSKLWPTTFFPLVKAFETKLKLSTLVLLHWPQGFSG 119

Qy 125 GQGPBK----IGPDGKXVILKDLTENPPTWRAMEKIVYEDRKAESQGSNWTIADLEK 178
 Db 120 DDLFPKDKGNAIG--GKATFL----DANEAMEELVDEGLVYALGVSNFSHQIEK 169

Qy 179 M-SKEPKVMPHANOIEHPFLPNELLVOYCFSKIMPVASPLGSQNQVPTTGERVS-- 234
 Db 170 LANKPGKXKPYTNQVECHPYLTQEKLQYCHSKGTTVATSPLGSPDRWAKPDPSSL 229

Qy 225 ENKTLNEAEGNTNLAQVLIAWGLRGGVYLPKSSNPKRIESNFKSIELSDADFEAINA 294
 Db 230 EDPKIKEIAAGHKKTAQVLIRFHIOVNVYIPKSVTPARIVENQVDFKLSDEMATI 289

Qy 295 VAKGRHFRFVNMKDITFGYDWP 316
 Db 290 LSFNRNRACNVLQSSHLDYP 311

Qy RESULT 15
 Db US-09-614-124B-340

Qy Sequence 340; Application US/09614124B
 / Patent No. 6630574
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Tongtong
 / APPLICANT: Bangur, Chaitanya S.
 / APPLICANT: Lodes, Michael A.
 / APPLICANT: Fanger, Gary
 / APPLICANT: Vedwick, Tom
 / APPLICANT: Carter, Darrick
 / APPLICANT: Retter, Marc
 / APPLICANT: Mannion, Jane
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
 / TITLE OF INVENTION: DIAGNOSIS OF LONG CANCER
 / CURRENT APPLICATION NUMBER: US/09/614,124B
 / CURRENT FILING DATE: 2001-07-11
 / NUMBER OF SEQ ID NOS: 1668
 / SOFTWARE: FastSEQ for Windows Version 3.0
 / SEQ ID NO: 340
 / LENGTH: 316
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-614-124B-340

Query Match 28.1%; Score 487; DB 4; Length 316;
 Best Local Similarity 36.6%; Pred. No. 8.8e-44;
 Matches 112; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

Qy 6 TFT-LSGVVKIPGVGFTFASEGSKGEVTYAVTALKTGRLDCAWYVLINEGEVGIR 64
 Db 3 TFEVLSTKAKNPPIVGLGTWKS - PLGKVKEAVKVAIDAGTRHIDCAYVYNEHEYGEAQ 60

Qy 65 DFLKENPSVKREDIFVCTKVNHLHYEDVLSIDSLSLRGLDVTMFLVHMPAAEKN 124
 Db 61 EKIQEK-AVKREDLFTVSKLWPTTFFPLVKAFETKLKLSTLVLLHWPQGFSG 119

Qy 125 GQGPBK----IGPDGKXVILKDLTENPPTWRAMEKIVYEDRKAESQGSNWTIADLEK 178
 Db 120 DDLFPKDKGNAIG--GKATFL----DANEAMEELVDEGLVYALGVSNFSHQIEK 169

Qy 179 M-SKEPKVMPHANOIEHPFLPNELLVOYCFSKIMPVASPLGSQNQVPTTGERVS-- 234
 Db 170 LANKPGKXKPYTNQVECHPYLTQEKLQYCHSKGTTVATSPLGSPDRWAKPDPSSL 229

Qy 225 ENKTLNEAEGNTNLAQVLIAWGLRGGVYLPKSSNPKRIESNFKSIELSDADFEAINA 294
 Db 230 EDPKIKEIAAGHKKTAQVLIRFHIOVNVYIPKSVTPARIVENQVDFKLSDEMATI 289

Qy 295 VAKGRHFRFVNMKDITFGYDWP 316
 Db 290 LSFNRNRACNVLQSSHLDYP 311

Qy RESULT 14
 / Sequence 340; Application US/09736457
 / Patent No. 6569448
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Tongtong
 / APPLICANT: Bangur, Chaitanya S.
 / APPLICANT: Lodes, Michael A.
 / APPLICANT: Fanger, Gary
 / APPLICANT: Vedwick, Tom
 / APPLICANT: Carter, Darrick
 / APPLICANT: Retter, Marc
 / APPLICANT: Mannion, Jane
 / APPLICANT: Pan, Liqun
 / APPLICANT: Wang, Ajun
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 / TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 / FILE REFERENCE: 21:0121_478C15
 / CURRENT APPLICATION NUMBER: US/09/736,457
 / CURRENT FILING DATE: 2000-12-13
 / NUMBER OF SEQ ID NOS: 1864
 / SOFTWARE: FastSEQ for Windows Version 3.0
 / SEQ ID NO: 340
 / LENGTH: 16
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-736-457-340

Query Match 28.1%; Score 487; DB 4; Length 316;
 Best Local Similarity 36.6%; Pred. No. 8.8e-44;
 Matches 118; Conservative 58; Mismatches 122; Indels 24; Gaps 8;

Query Match 28.9%; Score 499.5; DB 4; Length 316;
 Best Local Similarity 39.4%; Pred. No. 4e-45; Mismatches 106; Indels 29; Gaps 10;
 Matches 121; Conservative 51; Gaps 10;

Db 225 KIGKYGKSAAQVIIIRWHLQNDIIVIPKSVHEKRIOENFAVEDFELTPEEITAINQLYKD 284
 Qy 9 LSNGVKISPGVGFCTFASEGSKGEBTYTAVTTALKTGYRHLDCAWYLYNEGEVGEGTRDFLK 68
 Db 7 LNNGTKNTHPLGLGTWKS - -PPQQTAEVKVAIDGMYRHDCAQVYQNEKEVGVVALQECLK 64
 RESULT 12 US-09-347-803-12
 ; Sequence 12, Application US/09347803
 ; Patent No. 6271379
 ; GENERAL INFORMATION:
 ; APPLICANT: Famodu, Layo O.
 ; ATTORNEY: Kiney, Tony
 ; APPLICANT: Oroczo, Buddy
 ; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
 ; FILE REFERENCE: BB-1176
 ; CURRENT APPLICATION NUMBER: US/09/347,803
 ; CURRENT FILING DATE: 1999-07-02
 ; EARLIER APPLICATION NUMBER: 60/052,952
 ; EARLIER FILING DATE: July 15, 1998
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 12
 ; LENGTH: 308
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; US-09-347-803-12

Query Match 28.6%; Score 494.5; DB 3; Length 308;
 Best Local Similarity 36.4%; Pred. No. 1.3e-44; Mismatches 123; Indels 15; Gaps 8;
 Matches 114; Conservative 61; Mismatches 123;

Qy 69 LSNGVKIPGVGFCTFASEGSKGEBTYTAVTTALKTGYRHLDCAWYLYNEGEVGEGTRDFLK 68
 Db 4 LSSGHMTMSVGJGVWRMDSP - AIRDLIHSAIRIGRFHFDADYNEAEYGDALEAF - 60
 RESULT 9 USNSVKBREDIFVCTKVNHLHYEDVILWSIDDSKLKGLDYDVMFLVHMPTAEEKNQGQ-
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; TITLE OF INVENTION: ENTEROCOCCUS FACALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: 032796-032
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SEQ ID NO 6059
 ; LENGTH: 298
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-134-000C-6059

Query Match 28.8%; Score 498; DB 4; Length 298;
 Best Local Similarity 38.9%; Pred. No. 5.2e-45; Mismatches 48; Indels 40; Gaps 7;
 Matches 118; Conservative 97; Gaps 97;

Db 23 MTLDKTKKLLANGNEMPLGLGWRVEDG-SHATNSVKAWEVGRLIDAAYKNEVGK 81
 RESULT 13 US-09-702-705-340
 ; Sequence 13, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongrong
 ; ATTORNEY: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedwick, Tom
 ; APPLICANT: Carter, Derrick
 ; APPLICANT: Rettet, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun

Qy 1 MSNGKFTPLSNGKVKPGVGFCTFASEGSKGEBTYTAVTTALKTGYRHLDCAWYLYNEGEVG 60
 Db 23 MTLDKTKKLLANGNEMPLGLGWRVEDG-SHATNSVKAWEVGRLIDAAYKNEVGK 81
 RESULT 14 US-09-702-705-341
 ; Sequence 14, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongrong
 ; ATTORNEY: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedwick, Tom
 ; APPLICANT: Carter, Derrick
 ; APPLICANT: Rettet, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun

Qy 61 EGIRDFLKENPSVKREDIFVCTKVNHLHYEDVILWSIDDSKLKGLDYDVMFLVHMPA 120
 Db 82 EGIR---QSGIPREEIFTTKLYNEDOGENAIAKAFRSLSLEKLGTDYDLYIHWMPA 136
 Qy 121 AEKNGQGPKPGDGYKVILKDLTNPEPWRAKEVKYEDRKARSIGSVNTIADLEXMS 180
 Db 137 -----GRX-----KESWKALEBIIYASGRATAIGYNSFHQHLEDI 172
 Qy 181 KFAKYMMPHANOIEHPLPNBELVOYCFSKNIMPVAYSPLGSQNQVPTGERVEREENKTN 240
 Db 173 TEANVYPMQDQIELAPTLTQEPPLRKYLAENKIAEWAPSQQ-----GNLSE-PVLY 224
 Qy 241 BIAEKGGNITLAQVLIAWGLRGGYVYLPKSSNPKRITESNPK - -SIELSDADFEAINAVARG 298

Best Local Similarity 39.4%; Pred. No. 4e-45; Matches 121; Conserved 29; Gaps 10;

Qy 6 TFTLNSGVKIPGVGIGSTFAEBSGSGETTYAATTALKTGFRHLDCAWYTLNEGVGEGRD 65
Db 3 TVTLLGGYENHEVINGLWRRE-KDEBLKVLNIAKIGRHFDAAHYKSEADGEALAE 60
Qy 66 FLKENPSKREDIFVCTKWNHHLRYEDVWSDSLKLRLGLDYDMDPLVHWPIAAEKG 125
Db 61 AFKTG-LVKREBELLTKWNSDISH- -VWEACKNSLEKQIYDLTVLYHYPMPTEHNA 117
Qy 126 QGEPK- -IGPDGYVILKDLTENPPTWRAEMEKEYDRKARSIGVSNTIADLERMSKFA 183
Db 118 IGKTASLGED- -KVLDIDVTSISLQQTWEGMEXTVSLGLRSIGLSNVLFTRDCLAYS 175
Qy 184 KMPFHANQIEHPPFLNEEVQYCPSKINMPVATSPLGQNQVPTGEVS- -ENKLINE 241
Db 176 KIKPAVSQPETHPYFORDSLVKECMKHGVLPATAHTPLGGAANAKDMFGSSVPLDDPVIND 235
Qy 242 IAEGGGNTLIAQYLIAWGLRCCYVWLKPSSNPKRIESNFKIE--LSDADFEAINAVAKR 299
Db 236 VARKYGRSVAJICLAWGIQRTAVIPKSSKIQRLKENLVELEPQSLDEMQLYI- -DR 293
Qy 300 HFRFYNMKDTFGYDN 315
Db 294 KYRTSLPSKTIWGLDVY 309

Qy 9 LSNGVKIPGVGIGSTFAEBSGSGETTYAATTALKTGFRHLDCAWYTLNEGVGEGRD 68
Db 7 LNNGTKPTLGRGTWS- -PGQUTEAVKVADMYRHDCAQYQNEKEVGVALQEKLX 64

Qy 69 ENPSVREDFIDFCTKWNHHLRYEDVWSDSLKLRLGLDYDMDPLVHWPIAAEKG 128
Db 65 EQ-VVKGQDLTIVSKLWCTFDQSMVKGACQKTSQDLDYLDLYLHWWTGF----- 116

Qy 129 PKIGDPDKYVILKDLTENPPTWRAEMEKEYDRKARSIGVSNTIADLERMSKFA 179
Db 117 -KPGPD- -YFPL- DASGNVVISDTDFFVDTNAMEQIVDEGIVKAQGSNNPLOERILN 172
Qy 180 SKPAKYMHPHANOIEHPPFLNEEVQYCPSKINMPVATSPLGQNQVPTGEVS- -ENK 237
Db 173 KGLKXYPKAVNQIECHPLYIOEKLEYCHKGIVVTAQPSGPSPRWAKEPDPSLLEDP 232

Qy 238 TLINEIAFKGGNTLIAQYLIAWGLRCCYVWLKPSSNPKRIESNFKIE--LSDADFEAINAV 295
Db 233 RIKETIAKYNTKTAQVLIRFPIONLUVIPKSVTPARIAENFKVFDLSEMDATL- -L 290

Qy 296 AKGRHFR 302
Db 291 SYNARVR 297

RESULT 9
US-08-801-344-4
Sequence 4, Application US/08801344
Patent No. 6087140

GENERAL INFORMATION:
APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.
TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
TITLE OF INVENTION: FROM SUGAR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: DeWitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914
ZIP: 53717-1914
CITY: U.S.A.

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,344
FILING DATE: 03/25/2004
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 038820.037

TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Rat lens aldose reductase
ORGANISM: Rat lens aldose reductase

Query Match 28.9%; Score 499.5; DB 3; Length 316;

RESULT 10
US-09-498-599-4
Sequence 4, Application US/09498599
Patent No. 6303352

GENERAL INFORMATION:
APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.
TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL FROM SUGAR
TITLE OF INVENTION: 1,2-PROPANEDIOL FROM SUGAR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: DeWitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914
ZIP: 53717-1914
CITY: U.S.A.

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,599
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 038820.037

TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Rat lens aldose reductase
ORGANISM: Rat lens aldose reductase

US-09-498-599-4

RESULT 7
 US-08-731-320B-4
 Sequence 4, Application US/08731320B
 Patent No. 6116985
 GENERAL INFORMATION:
 APPLICANT: Wayne D. Loescher,
 APPLICANT: John D. Everard,
 APPLICANT: Rebecca Grumet
 TITLE OF INVENTION: DNA Encoding Mannose
 Phosphate Reductase
 TITLE OF INVENTION: 6-Phosphate Reductase
 TITLE OF INVENTION: Produced Therefrom
 NUMBER OF SEQUENCES: 4

ADDRESSEE: Ian C. McLeod
 STREET: 2190 Commons Parkway
 CITY: Okemos
 STATE: Michigan
 COUNTRY: USA
 ZIP: 48864

COMPUTER READABLE FORM:
 MEDIUM TYPE: Disquette 5.25 inch,
 MEDIUM TYPE: 360 kb storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 OPERATING SYSTEM: (version 3.3)
 SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/731,320B
 FILING DATE: October 15, 1996
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Ian C. McLeod
 REGISTRATION NUMBER: 20,931
 REFERENCE/DOCKET NUMBER: NSU 4.1-275
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (517) 347-4100
 TELEFAX: (517) 347-4103
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 310
 TYPE: Amino Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 MOLECULE TYPE:
 DESCRIPTION:
 HYPOTHETICAL: No
 ANTI-SENSE: No
 ORIGINAL SOURCE:
 ORGANISM: apple
 IMMEDIATE SOURCE:
 LIBRARY: N/A

US-08-731-320B-4

Query Match 29.0%; Score 501.5; DB 4; Length 310;
 Best Local Similarity 34.8%; Pred. No. 2.3e-45; Mismatches 73; Indels 15; Gaps 8;
 Matches 110;

Qy 6 TFTLNSGKVKGPGFGEASEGSKETTAYTTALKTGHRHLDCAWYVLYNEGYEGEGRD 65
 Db 3 TTVTLSSSYEMPVIGLGLWRL--KDELKEYTILNAKIGTRHFDCAHYSEADGEALE 60

Qy 66 FLKENPSVKREDIFVCTKWNHLHRYEDVLSIDSISLKLIGLDYDMFLYHPIAAEKG 125
 Db 61 AFKTG-LVKRBEFLITTKWNSDHGH--VTEACKNSLKLQDLYVHYPPTKNA 117

Qy 126 QGPBK--IGPDKYVILKDLTENPBTWAMEKIVEDRARSIGSVNTIALEKMSKA 183
 Db 118 IGBTASLGED--KVLDIDVTISLQQTWECEMKTYSLGIVRSIGSNYELFLTRCLAYS 175

Query Match 29.0%; Score 501.5; DB 4; Length 310;
 Best Local Similarity 34.8%; Pred. No. 2.3e-45; Mismatches 73; Indels 15; Gaps 8;

RESULT 5
US-09-166-412-4
; Sequence 4, Application US/09166412
; Patent No. 6133504
; GENERAL INFORMATION:
; APPLICANT: Wayne D. Loescher,
; APPLICANT: John D. Everard
; APPLICANT: Rebecca Grumet
; TITLE OF INVENTION: DNA Encoding Mannose
; 6-Phosphate Reductase
; TITLE OF INVENTION: and Recombinants
; TITLE OF INVENTION: Produced Therefrom
; NUMBER OF SEQUENCES: 4

Db 237 KYGKTPAQLVLRWGLQRDTTVIPKTSKVRLQENFDVFSPDISEDMERMKAI--DRKYR 294
Qy 303 FVNMDTGFQGDW 315
Db 295 TNQPAKFWGIDLY 307

Query Match 29.4%; Score 509.5%; DB 3; Length 309;
Best Local Similarity 37.6%; Pred. No. 3.2e-46;
Matches 118; Conservative 59; Mismatches 122; Indels 15; Gaps 8;

Db 4 TLNNQGKMPMLIGLGYWRMBG--EIRDLTINNSKIGYRFDCADYKNEAEVDALKEAF 61
Qy 8 TISNGYKIPGVGFPTFASSEGSKGETYTAVTALKTGYRHLDCAMYYLNEGGEVGEGIRD 67
Db 68 KENPSVKREDIFLVCWKWNAHRYEDVWLWDDLSIKRGLDYDMFLYWPPIAAEKGQG 127
Db 62 -DSGLVTKREDIFITPKLWNSSDOGH--VLEACKDSKLQLTYDLYHFPAVRHTGVG 118
Qy 128 EPK--IGPDGKVIILRDLTNEPEPTWRAMBKYYEDKARSIGSVNWTIADLERNSKPKV 185
Db 119 NTSSPQGDG--VLDDITTSLETWHAMELDLSSGLVSIGSINYDFILTDCLAYSKI 176
Qy 186 MPHANOIEIHPFLPNBELYCFSKRNIMPVSKPQVPTGERVS--ENKTLINEIA 243
Db 177 XPAVNQLETHPQFRSLVCPQRHICVTAHTPGAAANAEWFGTVSCLDDOVLKGIA 236
Qy 244 EKGGNTLAQVLIANGIIRGQYVLPKSSNPKRIESNFK--SIELSDADFFBAINAVAKGRHF 301
Db 237 EYKKTAAQSLRWRGIQQRNTVVPKSSKLRLKENFQVFDEELSKEDMELIGSI--DRKY 294
Qy 302 RFVNMDTGFQGDW 315
Db 295 RTNQPAKFWGIDLY 308

Query Match 29.4%; Score 501.3%; DB 3; Length 310;
Best Local Similarity 34.8%; Pred. No. 2.3e-45;
Matches 110; Conservative 73; Mismatches 118; Indels 15; Gaps 8;

Db 6 TFTLSNQVKIPGVGFPTFASSEGSKGETYTAVTALKTGYRHLDCAWYLYNEGEVGEGIRD 65
Qy 66 FLKENPSVKREDIFLVCWKWNAHRYEDVWLWDDLSIKRGLDYDMFLYWPPIAAEKG 125
Db 61 APKTG-LVKREBLFITTKINNSDGH--VVEACKNSLEXDQIYDLYLHVPMPTKNA 117
Qy 126 QEPBK--IGPDGKVIILRDLTNEPEPTWRAMBKYYEDKARSIGSVNWTIADLERNSKPKV 183
Db 118 IGKTASLGED--KVLDIDVTLISQQTWEGMEKTVSLGLRSIGHSYNEFLTRCLAY 175
Qy 184 KMMPHANQIEIHPFLPNBELYCFSKRNIMPVSKPQVPTGERVS--ENKTLINE 241
Db 176 KIKPAYSQFETHPQFRSLVCPQRHICVTAHTPGAAANAEWFGTVSCLDDOVLKGIA 235
Qy 242 TAEGGNTLAQVLIANGIIRGQYVLPKSSNPKRIESNFKLE--LSDADPAINAVAKGR 299
Db 236 VAKKYGKGSVAQICLRQIQQRKTAVIPKSSKLRLKENFQVFDEELSKEDMELIGSI--DR 293
Qy 300 HFRVNMDTGFQGDW 315
Db 294 KVTSSLSSKINGLDVY 309

Sequence 6, Application US/09347803
 General Information: Patent No. 674379
 Applicant: Famodu, Layo O.
 Applicant: Hitz, Bill
 Applicant: Kinney, Tony
 Applicant: Orozco, Buddy
 Title of Invention: Plant Sorbitol Biosynthetic Enzymes
 Title Reference: BB-1176
 Current Application Number: US/09/347,803
 Current Filing Date: 1999-07-02
 Earlier Application Number: 60/092,952
 Earlier Filing Date: July 15, 1998
 Number of SEQ ID Nos: 27
 Software: Microsoft Office 97
 SEQ ID NO 6
 LENGTH: 313
 TYPE: PRT
 ORGANISM: Glycine max
 US-09-347-803-6

Query Match 313; Score 543.5; DB 3; Length 313;
 Best Local Similarity 37.5%; Pred. No. 1e-47;
 Matches 123; Conservative 61; Mismatches 106; Indels 41; Gaps 10;

Qy 5 KTFPLSNGVKIPGVGFGTFASEGSKGETYTAVTTALKTGVRHLDCAWYVLINEGEVGEGIR 64
 Db 6 KFQLNLTGAKIPSVGLTWQAE--PGVVAKATAQVGYHIDCAAYKNQSEIGSALK 63

Qy 65 DFLKENPSVKREDIFVCTKWNHLHRYEDVLSIDSLLKRGLDYVDMFLVHWPIAEMK 124
 Db 64 K-LFDDGVVKREDLWTSKLWCSDESYDPKALKTLQDQLQDYLWVHWRMRSK 122

Qy 125 GOSEPCKIPGDGYVILDLTENPE--PTWAMEKUYEDRKAISGTVSNVTLIADEIKMSRF 182
 Db 123 SVG-----FKKEYLDQPDIBSTWMAEALYDSGARA.GVSNSSSKQLODANI 171

Qy 183 AKVMPHANQIEHPFLPNELVOYCFSKNIMPVAYSPLGSQNOVPTTGERSENKTNEI 242
 Db 172 ARVPANQVETLHPHQOQPKHAFPCSKVHLSGSPLGPGLVSKD--ILKNPVEIE 228

Qy 243 AEKGNTLAQVLIANGLRGYVVLPKSSNPKRIESNFKSIELSDAD----FEAI-NAV 295
 Db 229 AEKLGKTPAQALRNGLQTGHSPVNLKQDFDWSIPEVMWDKFSEIKQDR 288

Qy 296 AKGRHFRFVNMKDTEG---YDWPBEE 318
 Db 289 IKGTFF---VDETYGAFKTVHEELDGE 312

RESULT 3
 US-09-347-803-25
 Sequence 25, Application US/09347803
 General Information: Patent No. 674379
 Applicant: Famodu, Layo O.
 Applicant: Hitz, Bill
 Applicant: Kinney, Tony
 Applicant: Orozco, Buddy
 Title of Invention: Plant Sorbitol Biosynthetic Enzymes
 Title Reference: BB-1176
 Current Application Number: US/09/347,803
 Current Filing Date: 1999-07-02
 Earlier Application Number: 60/092,952
 Earlier Filing Date: July 15, 1998
 Number of SEQ ID Nos: 27
 Software: Microsoft Office 97
 SEQ ID NO 25
 LENGTH: 313
 TYPE: PRT
 ORGANISM: Medicago sativa
 US-09-347-803-25

Query Match 30.0%; Score 518.5; DB 3; Length 308;
 Best Local Similarity 37.7%; Pred. No. 3.4e-47;
 Matches 118; Conservative 60; Mismatches 120; Indels 15; Gaps 8;

Qy 9 LSNGVKIPGVGFGTFASEGSKGETYTAVTTALKTGVRHLDCAWYVLINEGEVGEGIRDFLK 68
 Db 4 LSSGHMRBAVGLGVWME-KADIRGLIHTLRLGYHRLCDAADQNEALVGDAEAFA- 60

Qy 69 ENPSVVKREDIFVCTKWNHLHRYEDVLSIDSLLKRGLDYVDMFLVHWPIAEEKGQG- 127
 Db 61 QTSLVKREDLFTTKLWMSDHG--VLEACKDSLKLQDYLWVHPPVATRHTGVGT 118

Qy 128 -ERKIGPDKVYTLKDNTPPTWAMEKUYEDRKAISGTVSNVTLIADEIKMSRFAKYM 186
 Db 119 TSSALGDG--VJDUDTISLETTWMAEELVSMGLRSIGSNTIPLTRDCLAWAKIK 176

Qy 187 PHANCIETHPELPNELVOYCFSKNIMPVAYSPLGSQNOVPTTGERBS--ENKTLNEIAE 244
 Db 177 PAYNQIETHPYFQRDSLVIKFCQKHGICVTAHTPLGGSTANAEWFGIVSCLDPVTSLAD 236

Qy 245 KGENTLAQVLIANGLRGYVVLPKSSNPKRIESNFK--SIELSDADFEA-NAVAKGRHFR 302

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 02:50:06 ; Search time 45 Seconds
(without alignments)
372.854 Million cell updates/sec

Title: US-10-004-115B-1
Perfect score: 1731
Sequence: 1 MSNGKFTLNSGVKPGVGF.....MKDTFGYDWPETAKNSA 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6C_POTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	547.5	31.6	344	4	US-09-530-583A-20	Sequence 20, Appl
2	543.5	31.4	313	3	US-09-347-003-6	Sequence 6, Appl
3	523.5	30.2	313	3	US-09-347-003-25	Sequence 25, Appl
4	518.5	30.0	308	3	US-09-347-003-10	Sequence 10, Appl
5	509.5	29.4	309	3	US-09-347-003-14	Sequence 14, Appl
6	501.5	29.0	310	3	US-09-166-412-4	Sequence 4, Appl
7	501.5	29.0	310	4	US-08-731-320B-4	Sequence 4, Appl
8	501.5	29.0	310	4	US-09-166-318-4	Sequence 4, Appl
9	499.5	28.9	316	3	US-08-801-344-4	Sequence 4, Appl
10	499.5	28.9	316	4	US-09-498-599-4	Sequence 4, Appl
11	494.8	28.8	298	4	US-09-134-000C-6059	Sequence 6059, Appl
12	494.5	28.6	308	3	US-09-347-003-12	Sequence 12, Appl
13	487	28.1	316	4	US-09-702-005-240	Sequence 340, Appl
14	487	28.1	316	4	US-09-736-457-340	Sequence 340, Appl
15	487	28.1	316	4	US-09-114-124B-340	Sequence 340, Appl
16	487	28.1	316	4	US-09-671-125-340	Sequence 340, Appl
17	487	28.1	316	4	US-09-599-184-340	Sequence 340, Appl
18	486.5	28.1	309	3	US-09-166-412-2	Sequence 2, Appl
19	486.5	28.1	309	3	US-09-347-003-26	Sequence 26, Appl
20	486.5	28.1	309	4	US-08-731-320B-2	Sequence 2, Appl
21	486.5	28.1	309	4	US-09-166-318-2	Sequence 2, Appl
22	472.5	27.9	308	3	US-09-347-003-8	Sequence 8, Appl
23	472.5	27.3	316	4	US-08-585-095-3	Sequence 3, Appl
24	466	26.9	291	4	US-09-107-532A-063	Sequence 503, Appl
25	456	26.3	323	4	US-09-702-005-1821	Sequence 1821, Appl
26	456	26.3	323	4	US-09-736-557-1821	Sequence 1821, Appl
27	456	26.3	323	4	US-09-671-125-1821	Sequence 1821, Appl

ALIGNMENTS

RESULT 1
US-09-530-583A-20
; Sequence 20, Application US/09630983A
; Parent No. 6630330
; GENERAL INFORMATION:
; APPLICANT: Porro, Danilo
; APPLICANT: Sauer, Michael
; TITLE OF INVENTION: Ascorbic Acid Production from Yeast
; FILE REFERENCE: 2028_594000
; CURRENT APPLICATION NUMBER: US/09/630,983A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 20
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-630-583A-20

Query Match 7 FTLSNGKTPKGVGFCFTAEGSKGETYATTAALKTGYRHLDCAWYYLNEGEVGEGRIRD 66
Best Local Similarity 31.6% ; Score 547.5; DB 4; Length 344;
Matches 122; Conservative 37.4%; Pred. No. 3.1e-50; Mismatches 58; Indels 55; Gaps 8;

QY 67 LKENPSKREDFLVCTKWNHLYRVEDILW----SIDDLSIKRGLDLYDUMFJWMPAA 121
Db 25 FSINNGNTRPAIGLGTANPHKELAFTKQAVRAAIKAGRHDIDATAVEYEPFVGAIKEL 84
QY 68 85 L-EDGSIKREDLFITFLTKVN-----PVLMDEVDSLNESKALGLEYVPLLLQHMPLCF 136
Db 67 LKENPSKREDFLVCTKWNHLYRVEDILW----SIDDLSIKRGLDLYDUMFJWMPAA 121
QY 88 122 EK-----NGQGPKTPGDPDKGKVILKDLTENPEPTGRAMEKLY--EDRKA 163
Db 137 EKIKDPKGISGIVKVKTPYDDSKTMTyAADGyL-----EYKQLEKITYLDNDHRY 186
QY 164 RSTIGVSNWTTABLEKMSKFAKVMPHANQIEIHPFLPNELVQYCFSKINMVPAYSPLGSQ 223
Db 87 RA/GVSNFSIYEYLRLIKECRVKPVTNQVETHPHLPQMELRKFCYMDILLTATSPGSH 246
QY 224 N---QVPTTGERVSSENKTNEIAEKGNITAQVLIJAGLRRGTYVLPKSSNPKRIESNFP 279
Db 247 GAPNLKP-----LVKLAERYNTGNLTLISYHQRTGTVIPRSUNPVRISSSI 296
QY 280 XSIELSDADFEAINAVAKGRHRFRYN 305
Db 247 GAPNLKP-----LVKLAERYNTGNLTLISYHQRTGTVIPRSUNPVRISSSI 296
QY 297 EPASLTDEOLQINDPQEKEYVRFID 322


```

; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS : 47374
; SEQ ID NO 1473
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-49-1473

Query Match Score 547.5; DB 15; Length 344;
Best Local Similarity 37.4%; Pred. No. 1.3e-47;
Matches 122; Conservative 58; Mismatches 91; Indels 55; Gaps 8;

Qy    7 FTLSNGVUKIPGVGEGTFASEBSKGETTAVTALKTYRHLDCAWYYLINEGEVGEGIRD 66
Db     25 FSUUNGVRIPALGLGTANPHXKLAETKQAVKAIAKAGCYRHDTANAYETEPVGAEIKEL 84
Qy    67 LKENPSVKREDIFVCTKWNHLHRYEDVLUW----SDDDSLRIGLDYVDFLVRPIAA 121
Db     85 L-EDGSIRKEDDLFITTWKW-----PVLNDEVDRSLNESLKAGLEYVDLLOHRPLCF 136
Qy   122 EK-----NGQGEERKIGPDGXVYLKDITENPPTRAMEKY---EDRKA 163
Db    137 EKIKDPKGISGLYVKTVPDDSGKTMYAADGDL-----ETYKOLEK2YLDPNIDHRV 186
Qy   164 RSIGVSNNTIADLEKMSKFAKMPHANOIEHFFLPNEELYQYCFSKNIMPVAYSPLGSQ 223
Db    187 RAIGVSNNSIEYERLIERLIKCRVPTVNQVETHHPOMELRKFCFNUFDILLTASPLGSH 246
Qy   224 N----QVPTIGERVSENKTLNEIAEKGENTLQQLIAGLRRGGYVLPKSNNPKRIESNF 279
Db    247 GAPNLKIP-----LVKLAEKAVTGNDLISYHROSTIVPRSLNPVRSSSI 296
Qy   280 KSTBLSDADFEANAVAKGRHFFEVN 305
Db    297 EFASLTKEDELQELNDFGERKYPVRFD 322

```

Search completed: May 28, 2004, 14:07:19
 Job time : 49 secs

RESULT 11
US-0-369-493-1683

Query Match Score 555; DB 15; Length 312;
Best Local Similarity 40.1%; Pred. No. 1..9e-48;
Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;

QY 172 TIADEKM-SKFAKMPHANOIEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 229
DB 172 SINNIKELLESPNKKVVPATNQTEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 230

QY 230 GERVSENKTLINEAEGNTNLQAQVLINGLRRGTVLPKSNSPKRISNFKSIELSDAF 289
DB 231 KE----QAIIDMAKXHGVEPPQLIISSWISQRGYVVLAKSNPERIVSNFKIFTLPEDDF 285

QY 290 EAINAVAKGRHF-RFVNMK 307
DB 286 KTISNLSKVHGTRKVUDMK 304

QY 2 SNGKFTLNSGVKIPGVGFTFASEGSKGETVAVTALKTYRHLDCAWYLNNEEVGE 61
DB 8 NSSATLKLITGASPVLGFTWSVDNNG-YHSVIAALKAGRHITAAATYLNEEEVGR 65

QY 62 GIRDFLKENPSVKREDIFVCTKVNHLHYEDLWSDLSLRGLDLYDMFLYHWPDA 121
DB 66 AIKD----SGVPRREEIFTIKLWGTEOR-DPEAALNKSLRGLDLYDMFLYHWPDA 118

QY 122 EKNGQGEPKKGPDGKYVILKDITE-----NPETWTAMEKTYEDRKARSIGVSNW 171
DB 119 KTDRV-----TDRNVLCIPTLEDGTVIDTKEWNFIKTTELWMOELPPTGKTKAVGVSNF 172

QY 172 TIADEKM-SKFAKMPHANOIEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 229
DB 173 SINNIKELLESPNKKVVPATNQTEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 229

QY 230 GERVSENKTLINEAEGNTNLQAQVLINGLRRGTVLPKSNSPKRISNFKSIELSDAF 289
DB 232 KE----QAIIDMAKXHGVEPAQLIISNSIQRGTVVLLAKSNPERIVSNFKIFTLPEDDF 286

QY 2 SNGKFTLNSGVKIPGVGFTFASEGSKGETVAVTALKTYRHLDCAWYLNNEEVGE 61
DB 7 NSSATLKLITGASPVLGFTWSVDNNG-YHSVIAALKAGRHITAAATYLNEEEVGR 64

QY 62 GIRDFLKENPSVKREDIFVCTKVNHLHYEDLWSDLSLRGLDLYDMFLYHWPDA 121
DB 65 AIKD----SGVPRREEIFTIKLWGTEOR-DPEAALNKSLRGLDLYDMFLYHWPDA 117

QY 122 EKNGQGEPKKGPDGKYVILKDITE-----NPETWTAMEKTYEDRKARSIGVSNW 171
DB 118 KTDRV-----TDRNVLCIPTLEDGTVIDTKEWNFIKTTELWMOELPPTGKTKAVGVSNF 171

QY 172 TIADEKM-SKFAKMPHANOIEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 229
DB 173 SINNIKELLESPNKKVVPATNQTEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 230

QY 230 GERVSENKTLINEAEGNTNLQAQVLINGLRRGTVLPKSNSPKRISNFKSIELSDAF 289
DB 231 KE----QAIIDMAKXHGVEPAQLIISNSIQRGTVVLLAKSNPERIVSNFKIFTLPEDDF 285

QY 290 EAINAVAKGRHF-RFVNMK 307
DB 286 KTISNLSKVHGTRKVUDMK 304

QY 2 SNGKFTLNSGVKIPGVGFTFASEGSKGETVAVTALKTYRHLDCAWYLNNEEVGE 61
DB 7 NSSATLKLITGASPVLGFTWSVDNNG-YHSVIAALKAGRHITAAATYLNEEEVGR 64

QY 62 GIRDFLKENPSVKREDIFVCTKVNHLHYEDLWSDLSLRGLDLYDMFLYHWPDA 121
DB 65 AIKD----SGVPRREEIFTIKLWGTEOR-DPEAALNKSLRGLDLYDMFLYHWPDA 117

QY 122 EKNGQGEPKKGPDGKYVILKDITE-----NPETWTAMEKTYEDRKARSIGVSNW 171
DB 118 KTDRV-----TDRNVLCIPTLEDGTVIDTKEWNFIKTTELWMOELPPTGKTKAVGVSNF 171

QY 172 TIADEKM-SKFAKMPHANOIEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 229
DB 173 SINNIKELLESPNKKVVPATNQTEIHPPLPNEELVQYCFSTKVNMPVAYSPIGSONQVPTT 230

QY 230 GERVSENKTLINEAEGNTNLQAQVLINGLRRGTVLPKSNSPKRISNFKSIELSDAF 289
DB 231 KE----QAIIDMAKXHGVEPAQLIISNSIQRGTVVLLAKSNPERIVSNFKIFTLPEDDF 285

QY 290 EAINAVAKGRHF-RFVNMK 307
DB 286 KTISNLSKVHGTRKVUDMK 304

RESULT 11
US-0-734-237B-39

Query Match Score 553; DB 15; Length 327;
Best Local Similarity 38.1%; Pred. No. 3..2e-48;
Matches 117; Conservative 69; Mismatches 105; Indels 16; Gaps 6;

QY Sequence 39 Application US/09734237B
DB Publication No. US2003006432A1
GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David

CURRENT FILING DATE: 2000-12-08
 PRIORITY APPLICATION NUMBER: 09/494,921
 PRIOR FILING DATE: 2000-01-31
 NUMBER OF SEQ ID NOS: 79
 SEQ ID NO: 44
 LENGTH: 313
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic protein derived from *Saccharomyces cerevisiae* Gcy1; reduced local similarity; conservative sequence
 OTHER INFORMATION: ctsase, having a glycine inserted at position 2 in the amino acid sequence
 US-03-734-237B-44

Query Match 33.2%; Score 574.5; DB 12; Length 313;
 Best Local Similarity 40.8%; Pred. No. 1.8e-50; Mismatches 54; Indels 45; Gaps 7;
 Matches 127; Conservative 54; N mismatches 85; Endata 11; AYRNEDQVGAIK 68

Qy 5 KTFSLNSNGKIPGVGFGTFASEGSKGETTAVTTALKTGIRHLDCAWYLNQEVGBGIR 64
 Db 11 KILSNTGAAQPQIGITWQSKE-ENDATAKVLALKDGYRHDAAIYRNEDQVGAIK 68

Qy 65 DFLKENPSVERDEFIVCTKWNHLHRYEDVLWDDSLXRLGLDYLDMPLVHMPIAEKN 124
 Db 69 D-----SGVREEIFVTKLWCTQHPEPV-ALDQSLXRLGLDYLDMMPARL-- 118

Qy 125 GQGPXKGPDGKVILKDLTENPE -----PTWRAFEMKYEDETRKASIGV 168
 Db 119 -----DPAYINNEIIISVPTKDKGSRAVDITNNNTIKIWLMQEPKIGTKRAYGV 169

Qy 169 SNWTIALEKRM--SKFKAMPHANOIEHTPFLPNEELYOCFSERNIMPVAYSPLGSONOV 226
 Db 170 SNSFINNLDDIASQNKLPAANGVEIPLLPODELINFCKSKGIVVVEAISPLGS--- 225

Qy 227 PTTGERVSERKTLNIAEKGNTLAQV1IAGLARGYVLPKSNSNPRIESNEKSIELSD 286
 Db 226 --TDAPBLRSPVILEAKNNVQPGHVVTSWHQRGVYVLPKSNSNPDRKTNRKIFTLS 283

Qy 287 ADFEAINAVAK 297
 Db 284 EDFEAINNISK 294

RESULT 8
 US-10-369-493-2285
 Sequence 2285, Application US/10369493
 Publication No. US200102333675A1
 GENERAL INFORMATION:
 APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Chen, Xianfeng
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 FILE REFERENCE: 38-10152054.B
 CURRENT APPLICATION NUMBER: US/10-369,493
 CURRENT FILING DATE: 2000-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374
 SEQ ID NO: 2285
 LENGTH: 321
 TYPE: PRT
 ORGANISM: *Schizosaccharomyces pombe*
 US-10-369-493-2285

Query Match 32.8%; Score 568; DB 15; Length 321;
 Best Local Similarity 43.0%; Pred. No. 8.8e-50; Mismatches 59; Indels 26; Gaps 10;

Qy 7 FTLSNGVKIFGVGFOTFASEGSKGETTAVTTALKTGIRHLDCAWYLNQEVGBGIRD 66
 Db 118 KTDRV----TDGNVLCIPTLEDGTVIDTKEWNFIKTELMOPLPKTGKTRAVGSNF 171

Db 16 FTIADSKIPGLGLGWRSE- PNQTKNAVXTALQYGYRHDAALYGNEDDEVGDIKE- 72
 Qy 67 LKENPSYKREDIFVCTKWNHLHRYEDVLWDDSLXRLGLDYLDMPLVHMPIAEKN 126
 Db 73 ---SGVPRKDIIWTSKLWCNAHAPPAVPALEKTIKDLCKDYLDLWIPVSF-KTGE 127
 Qy 127 GEPKIGPDGKVILKDLTENP-EPTWRAFEMKYEDETRKASIGVSNWTIALEKMSKFAKV 185
 Db 128 DKFPKDQDGNLIYEK---NPJETWRAFEMKYLETRKASIGVSNWTIALEKMSKFAKV 183
 Qy 186 MPHANQIEITHPFLPNEELYOCFSKNIMPVAYSPLGSON- -QVPTTGERVSERKNTLN 240
 Db 184 KPAVHQNEHLHPFLPQTEFVERHKKLGJHVTAYSPRONQNTIYESKIP---KLTHETIQ 239
 Qy 241 EIAE-XG-GNTLQAVILQJAWGURGTVLPISSNPKRIESNEKSIELSDADEFEAINAVAKG 298
 Db 240 KIAKSKSGBEGVGTGATIAWSWLTGTSVIPSNEQRIKSNSFYKIPLIKEDMDENSII-G 297
 Qy 299 RHERF 303
 Db 298 IARAF 302

RESULT 9
 US-09-801-368-438
 Sequence 438, Application US/09801368
 GENERAL INFORMATION:
 APPLICANT: Busby, Robert
 APPLICANT: Cali, Brian
 APPLICANT: Hecht, Peter
 APPLICANT: Holtzman, Doug
 APPLICANT: Madden, Kevin
 APPLICANT: Maxon, Mary
 APPLICANT: Milne, Todd
 APPLICANT: Royer, John
 APPLICANT: Salama, Sofie
 APPLICANT: Sherman, Amir
 APPLICANT: Silvaa, Jeff
 APPLICANT: Summers, Eric
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 FILE REFERENCE: 109272-147
 CURRENT APPLICATION NUMBER: US/09/801,368
 PRIORITY APPLICATION NUMBER: US 0020128250A1
 PRIORITY FILING DATE: 2001-03-07
 PRIORITY APPLICATION NUMBER: US 09/487,558
 PRIORITY FILING DATE: 2000-01-19
 PRIORITY APPLICATION NUMBER: US 60/160,587
 PRIORITY FILING DATE: 1999-10-20
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: Patentin version 3.0
 SEQ ID NO 438
 LENGTH: 312
 TYPE: PRT
 ORGANISM: *Saccharomyces cerevisiae*
 US-09-801-368-438

Query Match 32.1%; Score 555; DB 9; Length 312;
 Best Local Similarity 40.1%; Pred. No. 9e-48; Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;

Qy 2 SNGKTFIISNGVKIPGVGFOTFASEGSKGETTAVTTALKTGIRHLDCAWYLNQEVGBGVE 61
 Db 7 NSSATLKLNTGASIPVLFGTWRSVDNN--YHSVTAALKAGYRHDAAIYLNEEVEGR 64
 Qy 62 GIRDFLKRENPSYKREDIFVCTKWNHLHRYEDVLWDDSLXRLGLDYLDMPLVHMPIAA 121
 Db 65 AIKD---SGVPRKDIIWTSKLWCNAHAPPAVPALEKTIKDLCKDYLDLWIPVSF 117
 Qy 122 EKNGQCEPKIGPDGKVILKDLTENP-EPTWRAFEMKYEDETRKASIGVSNWTIALEKMSKFAKV 171
 Db 118 KTDRV----TDGNVLCIPTLEDGTVIDTKEWNFIKTELMOPLPKTGKTRAVGSNF 171

Query Match Score 972.5; DB 15; Length 255;
 Best Local Similarity 68.2%; Pred. No. 8.e-92; Indels 1;
 Matches 174; Conservative 35; Mismatches 45; Gaps 1;

Query 16 PGVGGTFASEGSKGETTAVATALKYRHLDCAWYLNSEBVEGIRDIFLKENNSVCR 75
 Database 1 PAVFGTFASEGSKGETTAAVKALEGVYRHLDCAWYLNSEBVEGQALAEFLNHDKVR 60

Query 76 EDIFCTKWNHHLERYEDLWSIDDSLRGLDGYVDMFLVHMPIAEKN-GGEPKIGPD 134
 Database 61 EDIFCTKWNHHLIPEDTKWSIDSLKDYVDEFLHMPIAEKDTEAMPKIGPD 120

Query 135 GKYVLLKDLTENPEPTWAKMEKLYEDRKARSIGSVNTIADLEKMSKPAKYMPHANOIEI 194
 Database 121 GKYLKKELLENPPPTWAKMEDLYDAGKTRSIGSVNTIPLQKLXPARKAKPTVNDIEI 180

Query 195 HPELPNNEELVOYCFSKNIMPVAVSPLGSQNQVPTTGERVSNTLNIAEKGGNTLAQVL 254
 Database 181 HPELPNTEELVEFCKQNQIPTAISPLGSQNQVPTTGERVDDTLKVAERGHNLAQVL 240

Query 255 IAWGLRRGTYVLPKS 269
 Database 241 IAWGLRRGTYVLPKS 255

RESULT 5
 US-09-734-237B-42
 ; Sequence 42, Application US/09734237B
 ; Publication No. US20030064432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bui, Peter
 ; APPLICANT: Hua, Ling
 ; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
 ; FILE REFERENCE: B583 : 40608
 ; CURRENT APPLICATION NUMBER: US/09/734, 237B
 ; CURRENT FILING DATE: 2000-12-08
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 42
 ; LENGTH: 312
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 ; US-09-734-237B-42

Query Match Score 574.5; DB 15; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.e-50;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

Query 5 KTFITLSNGVKGVPGFPTASEGSKGETTAVATALKYRHLDCAWYLNSEBVEGEGIR 64
 Database 10 KILSINTGAQPQIGLTVQSK--ENDAYKAVLTALKYRHLDCAWYLNSEBVEGEGIR 67

Query 65 DFLKENPSVREDIFVCTKWNHHLERYEDLWSIDDSLRGLDGYVDMFLVHMPIAEKN 124
 Database 68 D----SGVPRREITVTKWCTQHHEPEV--ALDQSXKRGLDGYVDMFLVHMPARL--- 117

Query 125 GQGEPKIGPDKYVILKDLTENPE-----
 Database 118 -----DPAYIKNEDILSVPTKDGSAVIDTNWNIPIKTWLMQELPTGKTKAVGV 168

Query 169 SNWTIADLEKM--SKFKAMPHANOIEHPLPNEEELVOYCFSKNIMPVAVSPLGSQNQV 226
 Database 169 SNSFSINNLKDILASOQNLTPAANQEVTHLPPQELINFCKSXGIVVEAYSPLGS--- 224

RESULT 6
 US-10-369-493-22372
 ; Sequence 22372, Application US/10369493
 ; Publication No. US2003023375A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10 (52052) B
 ; CURRENT APPLICATION NUMBER: US/10/369, 493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360, 039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO: 22372
 ; LENGTH: 312
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 ; US-10-369-493-22372

Query Match Score 574.5; DB 15; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.e-50;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

Query 5 KTFITLSNGVKGVPGFPTASEGSKGETTAVATALKYRHLDCAWYLNSEBVEGEGIR 64
 Database 10 KILSINTGAQPQIGLTVQSK--ENDAYKAVLTALKYRHLDCAWYLNSEBVEGEGIR 67

Query 65 DFLKENPSVREDIFVCTKWNHHLERYEDLWSIDDSLRGLDGYVDMFLVHMPIAEKN 124
 Database 68 D----SGVPRREITVTKWCTQHHEPEV--ALDQSXKRGLDGYVDMFLVHMPARL--- 117

Query 125 GQGEPKIGPDKYVILKDLTENPE-----
 Database 118 -----DPAYIKNEDILSVPTKDGSAVIDTNWNIPIKTWLMQELPTGKTKAVGV 168

Query 169 SNWTIADLEKM--SKFKAMPHANOIEHPLPNEEELVOYCFSKNIMPVAVSPLGSQNQV 226
 Database 169 SNSFSINNLKDILASOQNLTPAANQEVTHLPPQELINFCKSXGIVVEAYSPLGS--- 224

RESULT 7
 US-09-734-237B-44
 ; Sequence 44, Application US/09734237B
 ; Publication No. US20030064432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bui, Peter
 ; APPLICANT: Hua, Ling
 ; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
 ; FILE REFERENCE: B583 : 0608
 ; CURRENT APPLICATION NUMBER: US/09/734, 237B

Query Match Score 574.5; DB 15; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.e-50;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

Query 5 KTFITLSNGVKGVPGFPTASEGSKGETTAVATALKYRHLDCAWYLNSEBVEGEGIR 64
 Database 10 KILSINTGAQPQIGLTVQSK--ENDAYKAVLTALKYRHLDCAWYLNSEBVEGEGIR 67

Query 65 DFLKENPSVREDIFVCTKWNHHLERYEDLWSIDDSLRGLDGYVDMFLVHMPIAEKN 124
 Database 68 D----SGVREEFVTALKWCTQHHEPEV--ALDQSXKRGLDGYVDMFLVHMPARL--- 117

Query 125 GQGEPKIGPDKYVILKDLTENPE-----
 Database 118 -----DPAYIKNEDILSVPTKDGSAVIDTNWNIPIKTWLMQELPTGKTKAVGV 168

Query 169 SNWTIADLEKM--SKFKAMPHANOIEHPLPNEEELVOYCFSKNIMPVAVSPLGSQNQV 226
 Database 169 SNSFSINNLKDILASOQNLTPAANQEVTHLPPQELINFCKSXGIVVEAYSPLGS--- 224

Qy 61 EGIRDFLKENPSVKREDIFCYTKWNHLRYEDVLWSIDSLKRLGLDLYVDMFLVHPIA 120
Db 61 EGIRDFLKENPSVKREDIFCYTKWNHLRYEDVLWSIDSLKRLGLDLYVDMFLVHPIA 120

Qy 121 AERNGOGEPKIGPDKYVILKDLTENPEPTWRAKEV1YEDRKAISGVSNTIADLKMS 180
Db 121 AERNGOGEPKIGPDKYVILKDLTENPEPTWRAKEV1YEDRKAISGVSNTIADLKMS 180

Qy 181 KFAKMPHANOIEIHPFLPNEELYQCFSKNIMPVAYSPLGSONQVPTGERVSENKLN 240
Db 181 KFAKMPHANOIEIHPFLPNEELYQCFSKNIMPVAYSPLGSONQVPTGERVSENKLN 240

Qy 241 EIAEKGNTLAQVLIWAGLRRGGYVLPKSSNPKRIESNFSIELSDADEFEATNAVAKGRH 300
Db 241 EIAEKGNTLAQVLIWAGLRRGGYVLPKSSNPKRIESNFSIELSDADEFEATNAVAKGRH 300

Qy 301 FRFVNMDTFCYDVMPEETAKNLSA 325
Db 301 FRFVNMDTFCYDVMPEETAKNLSA 325

RESULT 2
US-10-327-108-3
; Sequence 3, Application US/10327108
; Publication No. US20030186400A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: WAKITA, RYUHEI
; APPLICANT: ITOH, NOUYA
; TITLE OF INVENTION: METHOD FOR PRODUCING OPTICALLY ACTIVE
; TITLE OF INVENTION: 2-HYDROXYCICLOALKANE CARBOXYLIC ACID ESTER
; FILE REFERENCE: Q73517
; CURRENT APPLICATION NUMBER: US/10/327,108
; CURRENT FILING DATE: 2003-12-24
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: JP 2001-395884
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: JP 2002-107648
; PRIOR FILING DATE: 2002-04-10
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 3
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Penicillium citrinum
; US-10-327-108-3

Query Match 100.0%; Score 1731; DB 14; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.5e-170; Mismatches 0; Indels 0; Gaps 0;

Db 1 MSNGKFTLNSGVKPGFPTASEGSKGETYAVTTALKTGRLDCAWYLINEGEVG 60
1 MSNGKFTLNSGVKPGFPTASEGSKGETYAVTTALKTGRLDCAWYLINEGEVG 60

Qy 1 MSNGKFTLNSGVKPGFPTASEGSKGETYAVTTALKTGRLDCAWYLINEGEVG 60
Db 1 EGIRDFLKENPSVKREDIFCYTKWNHLRYEDVLWSIDSLKRLGLDLYVDMFLVHPIA 120
1 EGIRDFLKENPSVKREDIFCYTKWNHLRYEDVLWSIDSLKRLGLDLYVDMFLVHPIA 120

Qy 121 AEKNGQEPKIGPDKYVILKDLTENPEPTWRAKEV1YEDRKAISGVSNTIADLKMS 180
Db 121 AEKNGQEPKIGPDKYVILKDLTENPEPTWRAKEV1YEDRKAISGVSNTIADLKMS 180

Qy 181 KPAKMPHANOIEIHPFLPNEELYQCFSKNIMPVAYSPLGSONQVPTGERVSENKLN 240
Db 181 KPAKMPHANOIEIHPFLPNEELYQCFSKNIMPVAYSPLGSONQVPTGERVSENKLN 240

Qy 241 EIAEKGNTLAQVLIWAGLRRGGYVLPKSSNPKRIESNFSIELSDADEFEATNAVAKGRH 300
Db 241 EIAEKGNTLAQVLIWAGLRRGGYVLPKSSNPKRIESNFSIELSDADEFEATNAVAKGRH 300

RESULT 3
US-10-608-533-1
; Sequence 1, Application US/10608533
; Publication No. US2004009192A1
; GENERAL INFORMATION:
; APPLICANT: ASAKO, HIROYUKI
; APPLICANT: SHIMIZU, MASATO SHI
; TITLE OF INVENTION: MODIFIED REDUCTASE AND ITS GENE
; FILE REFERENCE: Q76265
; CURRENT APPLICATION NUMBER: US/10/608,533
; CURRENT FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: JP 2002-193074
; PRIOR FILING DATE: 2002-07-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Penicillium citrinum
; US-10-608-533-1

Query Match 100.0%; Score 1731; DB 16; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.5e-170; Mismatches 0; Indels 0; Gaps 0;

Db 1 MSNGKFTLNSGVKPGFPTASEGSKGETYAVTTALKTGRLDCAWYLINEGEVG 60
1 MSNGKFTLNSGVKPGFPTASEGSKGETYAVTTALKTGRLDCAWYLINEGEVG 60

Qy 1 MSNGKFTLNSGVKPGFPTASEGSKGETYAVTTALKTGRLDCAWYLINEGEVG 60
Db 1 EGIRDFLKENPSVKREDIFCYTKWNHLRYEDVLWSIDSLKRLGLDLYVDMFLVHPIA 120
1 EGIRDFLKENPSVKREDIFCYTKWNHLRYEDVLWSIDSLKRLGLDLYVDMFLVHPIA 120

Qy 121 AEKNGQEPKIGPDKYVILKDLTENPEPTWRAKEV1YEDRKAISGVSNTIADLKMS 180
Db 121 AEKNGQEPKIGPDKYVILKDLTENPEPTWRAKEV1YEDRKAISGVSNTIADLKMS 180

Qy 181 KPAKMPHANOIEIHPFLPNEELYQCFSKNIMPVAYSPLGSONQVPTGERVSENKLN 240
Db 181 KPAKMPHANOIEIHPFLPNEELYQCFSKNIMPVAYSPLGSONQVPTGERVSENKLN 240

Qy 241 EIAEKGNTLAQVLIWAGLRRGGYVLPKSSNPKRIESNFSIELSDADEFEATNAVAKGRH 300
Db 241 EIAEKGNTLAQVLIWAGLRRGGYVLPKSSNPKRIESNFSIELSDADEFEATNAVAKGRH 300

RESULT 4
US-10-369-493-3396
; Sequence 3396, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldinan, Barry S.
; APPLICANT: Chen, Xianteng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(5205)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

ALIGNMENTS

SUMMARIES						
Query	Score	Match Length	DB ID	Sequence No.	Start	End
1	1731	100.0	325	14	US-10-004-11A-1	Sequence 1, Average
2	1731	100.0	325	14	US-10-327B-10B-3	Sequence 1, Average
3	1731	100.0	325	16	US-10-608-533-1	Sequence 1, Average
4	972.5	56.2	255	15	US-10-369-493-3396	Sequence 3.96
5	574.5	33.2	312	12	US-09-734-237B-24	Sequence 42, Average
6	574.5	33.2	312	15	US-10-359-493-22372	Sequence 223.72
7	574.5	33.2	313	12	US-09-734-237B-44	Sequence 44, Average
8	568	32.8	321	15	US-10-369-493-2285	Sequence 228.5
9	555	32.1	312	9	US-09-801-368-438	Sequence 43.8, Average
10	555	32.1	312	15	US-10-369-493-1683	Sequence 168.3
11	555	32.1	313	12	US-09-734-237B-39	Sequence 39, Average
12	553	31.9	327	15	US-10-369-493-2759	Sequence 275.9
13	548.5	31.7	315	15	US-10-369-493-2292	Sequence 229.2
14	547.5	31.6	313	12	US-10-424-599-282510	Sequence 282.510
15	547.5	31.6	344	15	US-10-365-493-1473	Sequence 147.3

DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.
 DR PROSITE; PS0062; ALDOKETO_REDUCTASE_2; 1.
 DR PROSITE; PS0063; ALDOKETO_REDUCTASE_3; 1.
 SQ SEQUENCE 309 AA; 35010 MW; 240501704DE05DEF CRC64;
 Query Match 30.6%; Score 529.5; DB 10; Length 309;
 Best Local Similarity 40.2%; Pred. No. 4.5e-35;
 Matches 119; Conservative 54; Mismatches 110; Indels 13; Gaps 7;
 Qy 8 TLNSGVKPIEGFGEFASEGSKGETYATTALKGYRILDCAWYLYNFGEGTDFL 67
 Db 4 TLNSGVKPIEGFGEFASEGSKGETYATTALKGYRILDCAWYLYNFGEGTDFL 61
 Qy 68 KENPSVKREDIFVCTPKVNHLHRYEDVMSIDDSLRKIGLDYDMLVPAEKNGQG 127
 Db 62 RTG-LVKREDLFTKLWNSDGHG--VIAACKDLRKLQDYLFLYFPATKTFVG 118
 Qy 128 -EPKIGPDKGYVILKDLENPEPTWRAMEKIYEDRKARSIGNSNWTADLEKMSKAKV 185
 Db 119 TTDSALGDDG-VLDLDTTSLETWHDKLYSMGLYTRISGNSNYDVLTRCLAYSKI 176
 Qy 186 MPHANQIEHPFLPNEELVQYCFSKNIMPVAYSPLGSONQVPTGER--VSEN 236
 Db 177 KPAVNOQTETPYFQDSTSIVKFCRGHICVTAHTPLGGATANAEWFGTSCLDPVLDKA 236
 Qy 244 ERGGNTLAVQLIANGLRRGVVLPKSSNPKRIBENFK-SIBLSDADEPAINAVAK 297
 Db 237 EKYKTTVAQVYLRNQIQRKVVIKTSKPARLESNQVDFEFSKEDMEVIRSMER 292
 Search Completed: May 28, 2004, 14:05:46
 Job time : 47 secs

RESULT 15
 080945 PRELIMINARY; PRT; 290 AA.
 AC 080945; PRELIMINARY;
 DT 01-NOV-1998 (TREMBiReL 08, Created)
 DT 01-JUN-2003 (TREMBiReL 24, Last sequence update)
 DE Putative alcohol dehydrogenase.
 AT2G3770; OS Arabidopsis thaliana (Mouse-ear cress);
 OC Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;
 OC Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI TaxID=3702;
 RN [1];
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv, Columbia;
 MEDLINE=20083487; PubMed=10517197;
 RA Lin X., Kaul S., Rounseley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii Y., Mason T.M., Bowman C.L., Barnstead T.N.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Crasy T.H., Goodman H.M., Somerville C.R.,
 RA Copanhauser G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.,
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 RT Nature 402:761-763(1999).
 RN [2];
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv, Columbia;
 RA Lin X.; Submitted (MAR-2000) to the EMBL/GenBank/DDJB databases.
 RL DR EMBL: AC04684; AAC23647.1; -.
 DR PIR: T02543; T02543.
 DR HSSP; P14550; 2ALR.
 DR InterPro: IPR001395; Aldo/ket red.
 DR Pfam: PF00248; aldo_ket_red_1.
 DR PRINTS; PR00059; ALDKEFDTASE.
 DR PROSITE; PS00798; ALDOKETO_REDUCASE_1; 1.
 DR PROSITE; PS0062; ALDOKETO_REDUCASE_2; 1.
 DR PROSITE; PS0063; ALDOKETO_REDUCASE_3; 1.

RESULT 13			
121	- - - - - AEKNCQGEPKIGPDKYVILKDLTENPEPTWAMEKITYDRKARSIGVSNW	Q2GXW0	PRELIMINARY;
Dy	121 IKDGNLNLFQAEK - - - - - PDGSKDY - - - - -	Q8GXW0	PRT; 322 AA.
Db	172 TIADEKMR - SKPAKVMPHANQTEIHPFLPNELVOYCFSKNIMPVAYSPLGOSNOVPTT	Q2GXW0	
Dy	170 SVNKLQDIAAPTPKTKUTPVAVQYELPLPQEDIKFPGENGIVTEASPLGEN- - -	Q8GXW0	
Db	230 GERVSENKTLINEAEGGGNTLAAQVJANGLRGYVVLPKSNPKRIESNFKSIEBLSDADF	Q2GXW0	
Dy	225 -APLSDPVPQELAKANGYEAGHRVVISAVQKGTLPLRSVTPSRIVGNLKVLTLSDDV	Q8GXW0	
Db	290 EAANAVAKGR 299	Q2GXW0	
Dy	284 AKVDGLLKAK 293	Q2GXW0	

DX NCBI_TAXID=3702;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=cv. Columbia;
 CRC
 Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
 Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
 Hayashizaki Y., Shinozaki K.;
 "Arabidopsis thaliana full-length cDNA.";
 RT

Submitted (NOV-2002) to the BMBL/GenBank/DDJB databases.
 ENBL: AK118010 ; BAC426A3_1 ; - .
 InterPro: IPR00135 ; Aldo/ket-red.
 Pfam: PF00248 ; aldo ket red; 1 .
 PRINTS: PRO0069 ; ALDKERKDTTAE.
 ProDom: PDO02288 ; Aldo/ket red; 1 .
 PROSITE: PS00062 ; ALDOKETO_REDUCTASE_2 ; 1 .
 PDR : PS00263 ; ALDOKETO_REDUCTASE_3 ; 1 .
 PROTEIN: 322 AA ; 36581 MW ; 1PEBE7459700FA2 CRC64 ;
 SEQIDENT : - .

Qy	182 FAKYMPHANOQEIPFPLPNEELVQYCFSKNIIMPAYSPIGSONQVPTGERSENKTLINE
Ddb	179 FAELIPAVCOMEMPGWNRDILFECRKNEHHTAYSPGSOE --- GORDLIDHQTYDR
Qy	242 TAEKGNTLAQVLIAWGLRGYVFLPKSSNPKR1ESNFKSE - LSDADFEAINAVKGR
Ddb	235 IAKLNKTPQQLVWKMGLQRGTSVPSLNPERIKENIYFPDWVIEPEQDFQALNSITDQK
Qy	300 HEREVNMKDTE ----- GYDYMPEE 318

Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.

DR	EMBL; AE003544; AAF00339_2; -;	PROSITE; PS00063; ALDOKETO REDUCTASE_3; 1.
DR	XW Oxidoreductase; Complete Proteome.	
SEQUENCE	279 AA;	31836 MW; D3458455BF85CF CRC64;
SQ		
Query Match	30.6%; Score 530.5; DB 16; Length 279;	
Best Local Similarity	40.3%; Pred. No. 3.2e-35;	
Matches	120; Mismatches 56; Indels 37; Gaps 9;	
Qy	1 MSNGKFT-LSGVKP1 PGVGRGTPASEGSKGETTYTAVTTALKTGYRHLDCAWYLYNEGEV 59	
Db	1 MENLQSKTVLSNGVENVWFGVGVFKEGP-ELVERAKSAIKAGRSIDTAIYGNERAAV 59	
Qy	60 GEIGIRDFLKENSVKREDIFVCTKVNHLRYEDVWISLSSKLGLDVKDMFLVHWP1 119	
Db	60 GEIGIRAGI-ETRGISNEELFTSKWNADQGYEETIAAYEESLKKLKDLYLVHWPV 118	
Query Match	30.7%; Score 531.5; DB 5; Length 316;	
Best Local Similarity	38.6%; Pred. No. 3.2e-35;	
Matches	124; Conservative 60; Mismatches 108; Indels 29; Gaps 10;	
Qy	10 SNGVKIPGFGTFASEGGSKGETTYTAVTTALKTGYRHLDCAWYLYNEGEVGEGIRDPE 69	
Db	119 -----EKY--KD---TWRALEYKEKRVTAIGVSNNQIHLQDV 154	
Qy	10 NPSVREDFIVCTKVNHLRYEDVWISDSSLRLGIGYDMLVHFRPIAAKNGQCEP 129	
Db	120 AAEKGNGOEPKIGPDKYVILKDLTNPEPTWRAMEKIYEDRKARSIGVSNNWTLADLERM 179	
Qy	11 G-VVREDFLTTSKLNNTFRPLDKYVLLTSSLKXLDLYLHWN-GYKEGCDLF 124	
Db	121 -----EKGKTAQVIRWDLONGVITPKSTREHHLIANADVNFELTSDMEKIDAL 264	
Qy	130 KIGPDCK-----YVILKDLTENBPBTWAMEKIYEDRKARSIGVSNNWTLADLERKSF 182	
Db	125 PTDKGKTKLSPVDPVY-----DTYKAMEKLVBEGLVKSIGVSNNRQRQERVLEV 174	
Qy	133 AKVMPHQANQIEBHFPFLPNPBDVQYQPSKRNIMPVAYSPLPSQNOV-PITGERV-SENKTLN 240	
Db	134 ATIPPVNTQIECHPYITQKLIDFKSKDITATYSPSPLQSPNRPWAKAGDPVILEAKIK 234	
Qy	135 EIAERGGNTAQVILAWGLERGGYVLPKESNSPKRNEESNPK-SIELSDDAEEAINAV-AK 297	
Db	136 EIAAKKKKTPGQILRYYQVORANIVPKSYTKDRIESNFQVDFDFELTPPEEITIESFBCN 294	
Qy	138 GHEFRVNMKDTFGYDWPRE 318	
Db	139 G---RLVPLDNQYGHPHHPHPP 312	
RESULT 11		
Q81B04	PRELIMINARY; PRT; 279 AA.	
ID	Q9HGKX9	PRELIMINARY; PRT; 310 AA.
AC	Q9HGKX9;	
DT	01-MAR-2001 (TRIMBLrel. 16, Created)	
DB	01-MAR-2001 (TRIMBLrel. 16, Last sequence update)	
Qy	241 EIAERGGNTAQVILAWGLERGGYVLPKESNSPKRNEESNPK-SIELSDDAEEAINAV-AK 297	
Db	242 EIAAKKKKTPGQILRYYQVORANIVPKSYTKDRIESNFQVDFDFELTPPEEITIESFBCN 294	
Qy	248 GHEFRVNMKDTFGYDWPRE 318	
Db	249 G---RLVPLDNQYGHPHHPHPP 312	
SEQUENCE FROM N.A.		
RC	STRAIN=ATCC42981;	
RA	Iwaki T., Kurono S., Yokose Y., Kubota K., Tamai Y., Watanabe Y.;	
RT	"Cloning of glycerol-3-phosphate dehydrogenase genes (zrgcy1 and zrgcy2) from the yeast Zymosaccharomyces rouxi."	
DR	Zrgcy1 and glycerol dehydrogenase genes (zrgcy1 and zrgcy2) from the salt-tolerant yeast Zymosaccharomyces rouxi.";	
RT	salt-tolerant yeast Zymosaccharomyces rouxi.";	
RL	0:0/0(2001).	
DR	Glycerol dehydrogenase.	
DR	EMLB; AB047397; BAB1960-2. -	
DR	HSSP; P066312; 1HW6.	
DR	InterPro; IPR01395; Aldo/ket red.	
DR	Pfam; PF00448; aldo_ket_red; 1.	
DR	PRINTS; PR00069; ALDKETREDTASE.	
DR	Zrgcy1 and glycerol dehydrogenase genes (zrgcy1 and zrgcy2) from the yeast Zymosaccharomyces rouxi.";	
DR	Zrgcy2.	
DR	PROSITE; PS00062; ALDOKETO_REDUCTASE_1. 1.	
DR	PROSITE; PS00052; ALDOKETO_REDUCTASE_2. 1.	
DR	PROSITE; PS00053; ALDOKETO_REDUCTASE_3. 1.	
DR	SEQUENCE; 310 AA; 34106 MW; 80IDBD1424775C375 CRC64;	
Qy	6 TFTLNGVKINGVGFTEASEGSKGETTYTAVTTALKTGYRHLDCAWYLYNEGEV/EGIRD 65	
DR	EMBL; AE017009; AAPJ0328.1; -	
GO	GO:0016491; F:oxidoreductase activity; IEA.	
InterPro	IPR001395; Aldo_ket_red.	
PFam	PF00248;aldo_ket_red; 1.	
ProDom	PD000388; Aldo_ket_red; 1.	
DR	PROSITE; PS00798; ALDOKETO_REDUCtASE_1; 1.	
DR	PROSITE; PS00062; ALDKETREDTASE.	
DR	SEQUENCE; 310 AA; 34106 MW; 80IDBD1424775C375 CRC64;	
Qy	66 FLKENNSVKREDIFVCTKVNHLRYEDVWISLSSKLGLDVKDMFLVHWP1 120	
Db	68 -----SSVPRNEIFLTKLWCTQQRNPQE--ALDQSLORIGLDVYLTHWPVLRLEN 120	

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC Orygenales; mitosporic Onygenales; Coccidioides.
OX NCBI_TaxID=5501;
RN [1] SEQUENCE FROM N.A.
RP DEJagado N., Cole G.T., Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RA InterPro: IPI001305; Aldo-/ket_red.
DR PFAM: PF00248; aldo-/ket_red_1.
DR PRINTS: PR00069; ALDKEFRDTASB.
DR PRODom: PD000288; Aldo-/ket_red_1.
DR PROSITE: PS00798; ALDOKETO REDUCTASE_1; 1.
DR PROSITE; PS0062; ALDOKETO REDUCTASE_2; 1.
DR SEQUENCE; 314 AA; 09CD9B7734334134 CRC64;
SQ SEQUENCE; 314 AA; 35411 MW; "The genome sequence of *Drosophila melanogaster*.";
RN [2] "The genome sequence of *Drosophila melanogaster*.";
RN [2] RT Science 287:2185-2195 (2000).
RN [2] RP SEQUENCE FROM N.A.
RA Celinker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A., Rogers Y.,
RA Evans C.A., Gocayne J.D., Anatolides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Centor A., Champé P.M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Douq L.E., Doyle C., Drysek D., Farfan D.,
RA Ferriera S., Frise E., Gallo R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Neilson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Richards S., Scheeler F.,
RA Phourenavong S., Pittman G.S., Puris V.,
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Annotation of *Drosophila melanogaster* genome";
RN [3] RT Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3] RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Clamp M., Drysdale R., Emmert B., Garrison J.W., Celinker S.E.,
RA Krommiller B., Marshall B., Millburn G., Richter J.A., Harris N.,
RA Seale S.M.J., Smith E., Shu S., Smirnios F., Whittlefield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome";
RN [4] RT Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4] RP SEQUENCE FROM N.A.
RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RN [5] RP Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [5] RP SEQUENCE FROM N.A.
RA FlyBase;
RN [6] RP Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN [6] RP SEQUENCE FROM N.A.
RA STRAIN=Y;
RA STRAIN=M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champé M., Chavez C., Dorsett V., Drysek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Mirand A., Mungall C.J., Russo S.,
RA Patel S., Phourenavong S., Wan K., Lewis S.E., Rubin G.M.,
RA Celinker S.;

RESULT 10
Q9VTK9 PRELIMINARY; PRT; 316 AA.
ID Q9VTK9
AC Q9VTK9
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-JUN-2003 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE CG6084 protein (LD06395P).
DE CG6084
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila;
OX NCBI_TaxID=7227;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RN MEDLINE=20195006; PubMed=10731132;
RA Adams M., Celinker S.E., Gocayne J.D.,
RA Amanatides P.G., Scheerer S.E., Li P.W., Ashburner M., Henderson S.N.,
RA George R.A., Lewis S.E., Richards S., Ashburner M.,
RA Sutcliffe G.G., Wortham J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon C., Rogers Y.-H.C., Blazej R.G., Champé P.M., Preiffner B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Neilson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhambhani D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RL	Yeast 0:0-0(2001)	DR	PROSITE; PS00063; ALDOKETO_REDUCtASE_3 ; 1.
EMBL;	AB047336; BAB11959.1; - .	KW	Complete proteome.
DR	HSSP; P0632; 1.HW.	SEQUENCE	327 AA; 36014 MW; 4B9415D098A8892D CRC64 ;
DR	InterPro; IP0001395; Aldo/ket red.	Query Match	31.9%; Score 553; DB 16; Length 327;
DR	PFAM; PF00248; aldo/ket red.1.	Best Local Similarity	38.1%; Pred. No. 5.9e-37;
DR	PRINTS; PR00069; ALDKETREDTASE_1.	Matches	117; Conservative 69; Mismatches 105; Indels 16; Gaps 6;
DR	ProdDom; PD00288; ALDOKETO_REDUCtASE_1.	Qy	5 KTFITSGNGVKIPGYGFGTFASEGSKGGETTAVTALKGYRHDCAWYLYNNEGVGEIGIR 64
DR	PROSITE; PS000798; ALDOKETO_REDUCtASE_1.	Db	11 KYFPLNSGEQLPAQGLTWKS--SPQVWQAVEQALDGYRHDCAYGNNEAIGATIA 68
DR	PROSITE; PS00062; ALDOKETO_REDUCtASE_1.	Qy	65 DFLKENPSVKREDIFVCTKWNHLERYDVLWSDDSLKRGLGDYVMDFLVHPIAAEKN 124
DR	PROSITE; PS00063; ALDOKETO_REDUCtASE_1.	Db	69 NAFTKG-VVKREELWITSKWSNAHHHPDAVLPALEKTLDGLDYLILHNPVVI--- 124
SQ	SEQUENCE 310 AA; 34126 MW; F201FFBB8CCB9CFD CRC64 ;	Qy	125 QGGEPIKG-PDGKVTLKDLTENDEPTFAMEKLYEDRKARSIGVSNTIADLJKMSKA 183
DR	10 TLXLTNGQTTPQVGLGTWSKNEG--YKAVIBALKAGYRHTDGAAYGNBEGEVKAID 67	Db	125 --QPDVGFPESGQDOLPPFTPASLEGTVQALEXAVDGLCHHIGVSNSSLKLEMVLSMA 181
Qy	66 FLKPNPSVTKREDIFVCTKWW-NFLHRYDVLWIDDLSKRLGIDYTMFLVHPIA--- 120	Qy	184 KVMPHANQTHPPVNLNEVOYCFSKNIMPVASQVLLGRRGGYVVLPKSSNPKRIESNEKFSTELSDADFEAINAYKG 239
Db	68 ----SGVPRNEFLTTKLWCTHORNPOAL--DQSISQRGIDyVSYLHWPVPDRT 119	Db	182 RIPPVNQVLPHQPLQSDLTFTANSQNILLTAYSPLOSSGDR PAAPFOQAEPKLITDPV 240
Qy	121 -----AEKNGQGEPKIGPDGKYVILKOLTENPEPTFRAAMEKLYEDRKARSIGVSN 170	Qy	240 -NEIAEKGENTLAQVLIANGLRRGGYVVLPKSSNPKRIESNEKFSTELSDADFEAINAYKG 298
Db	120 NIKDGNLFQFAEK-----PDGSKDV--DLEWNFIKTWEMLQKLESGXTKAIVGSN 168	Db	241 INGIAEQGSSAQVLLAWAIQRGTVTTPKSVNBRLBONLRADITLTDSENAKIALD 300
Qy	171 WTIADELM-KSKPARKMHPHQNLIEHPPLNBEVQCFSKNIMPVASQVLLGRRGGYVVLPKSSNPKRIESNEKFSTELSDAD 228	Qy	299 RHEFVN 305
Db	169 FSVNNLKDLLAAPTAKTPKVTPAVNVQVEFHULLPQBELIQCSEKIVTEAYSTPLGEN--- 224	Db	301 RHYRYS 307
Qy	229 TGBERVSNTLNEAEGKNTNLQVLTQVLTAWGLRGGYVVLPKSSNPKRIESNEKFSTELSDAD 288	RESULT 6	
Db	225 -APILSDPTVQEAKANGVTAQGHVVISWAQKGLVTLPKSVTPS1RVGNLKVLTSDSD 282	ID	O42888 PRELIMINARY; PRT; 325 AA.
Qy	289 FEAINAVAKGR 299	AC	O42888; PRELIMINARY; PRT; 325 AA.
Db	283 VAKVDGLLRAK 293	DT	01-JUN-1998 (TREMBLrel. 06, Created)
Qy		DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)
Db		DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE	Probable oxidoreductase.	DE	SPBCB4.04.
GN	SR0942.	OS	Schizosaccharomyces pombe (Fission yeast).
OS	Synechocystis sp. (strain PCC 6803)	CC	Eukaryota; Fungi; Ascomycota; Schizosaccharomyces pombe.
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.	CC	Schizosaccharomyces; Schizosaccharomyces pombe.
RN	[1] NCBI_TaxID=1148;	NCBI	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RX	P74308	RC	STRAIN=972h-;
AC	P74308;	RC	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DT	01-FEB-1997 (TREMBLrel. 02, Created)	RA	Wood V., Barrell B.G., Rajandream M.A.;
DT	01-JUN-2003 (TREMBLrel. 02, Last sequence update)	RA	EMBL: AL021815; CAI61997.1; -
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	RA	PIR: T50378; T39169.
DE	Aldehyde reductase.	RA	HSSP; PI4150; 2ALR.
GN		RA	GeneDB; SPBCE8E4_04; -
OS	Synechocystis sp. (strain PCC 6803)	RA	InterPro; IPR001395; Aldo/ket red.
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.	RA	PF00248; aldo/ket red; 1.
RN	[1] NCBI_TaxID=1148;	RA	PRINTS; PRO00059; ALDKETREDTASE.
RP	SEQUENCE FROM N.A.	RA	ProdDom; PD00028B; Aldo/ket red; 1.
RX	P74308; PubMed=8905231;	RA	PROSITE; PS00193; ALDKETTO_REDUCtASE_1; 1.
RA	01-FEB-1997 (TREMBLrel. 02, Created)	RA	PROSITE; PS0062; ALDKETTO_REDUCtASE_2; 1.
RA	01-JUN-2003 (TREMBLrel. 02, Last sequence update)	RA	PROSITE; PS00063; ALDKETTO_REDUCtASE_3; FALSE NEG.
RA	Aldehyde reductase.	RA	SEQUENCE 325 AA; 36629 MW; 68ED21F1E4D08D3 CRC64 ;
RT	"Sequence analysis of the unicellular cyanobacterium	DR	Query Match 31.7%; Score 548.5; DB 3; Length 325;
RT	synechocystis sp. strain PCC6803. II. Sequence determination of the	DR	Best Local Similarity 40.5%; Pred. No. 1.4e-36;
RT	entire genome and assignment of potential protein-coding regions."	DR	Mismatches 52; Indels 37; Gaps 8;
RL	DNA Res. 3:109-136(1996);	DR	
DR	EMBL; D50914; BAA18402.1; - .	DR	
DR	PIR; S76143; S76143.	DR	
DR	HSSP; PI4450; 2ALR.	DR	
DR	InterPro; IPR00135; Aldo/ket red.	DR	
DR	PRINTS; PRO00059; Aldo/ket red.	DR	
DR	ProdDom; PR00069; ALDKETREDTASE.	DR	
DR	PROSITE; PS00288; Aldo/ket red; 1.	DR	
DR	PROSITE; PS00063; ALDKETTO_REDUCtASE_1; 1.	DR	
DR	InterPro; IPRO0135; Aldo/ket red.	DR	

Qy	121 AEKNGQEPKPGDKYVILKDLTENPEPTWAMEKIYEDRKARSIGVSNTIADLEKMS 180	DT 01-NOV-1998 (TREMBLrel. 08, Created)
Db	121 SEKDOEQPKPGDKYVILKDLTENPEPTWAMEKIYEDRKARSIGVSNTIADLEKMS 180	DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
Db	181 KFAKYMMPHANOIEHPFLPNEELYOYCFSKINMPVAYSPLGSONQVPTGERSENKLNN 240	DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
Qy	181 KFAEKPHINQIEHPFLPANELYQCFNDLPEAYSLGSQNQVPTGERSENKLNN 240	DE Reductase (Frasier).
Db	Hypocreomycetidae; Ascomycota; Pezizomycotina; Sordariomycetes;	OS Gibberella zeae (Fusarium graminearum).
Db	NCBI - TaxID:5518;	OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Qy	241 EIAEKGGNNTLAQVLIWGLRRGGYVLPKSSNPKRIESNPKSIELSDAEEANAVAKGRH 300	OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
Db	241 EIAQKGGNNTAQVLIWGLRRGGYVLPKSSNPARIESNPKSIELDEDAEVNKVAEGRH 300	OX NCBI - TaxID:5518;
Db	SEQUENCE FROM N.A.	RN
Db	STRAIN=F15;	RP
RC	MEDLINE=99433864; PubMed=9762900;	RC
Qy	"The mystery of the trichothecene 3-O-acetyltransferase gene. Analysis of the region around Tr101 and characterization of its homologue from Fusarium sporotrichioides."	RX Kimura M., Matsumoto M., Shiozaki Y., Yoneyama K., Yamaguchi I.; RT RT
Db	FEBS Lett. 435:163-168 (1998).	RT
EMBL	AB014493; BAA33773; 1;	RL
DR	PF00424; aldo_ket_red.	DR
DR	InterPro; IPR01395; Aldo_ket_red.	DR
DR	P06632; 1HW6.	DR
DR	PRINTS; PRO0069; AldKETTASE.	DR
DR	PRODOM; PD00288; Aldo_ket_red.	DR
DR	PROSTTB; PS00798; ALDOKETO_REDUCTASE_1; 1.	DR
FT	NON-TER 1	FT
FT	NON-TER 254	FT
SQ	SEQUENCE 254 AA; 28387 MW;	FT
Qy	Query Match 16	Query Match 16
Qy	Best Local Similarity 59.9%; Score 1037; DB 3; Length 254;	Best Local Similarity 59.9%; Score 1037; DB 3; Length 254;
Db	Matches 185; Conservative 30; Mismatches 39; Indels 0; Gaps 0;	Matches 185; Conservative 30; Mismatches 39; Indels 0; Gaps 0;
Qy	RT 1 PGVGFGTPEASBGSKGBFTYTAVTALKTGVRHLDCAWVYLNEGEVGEIRDFLKENPSVKR 75	RT 1 PGVGFGTPEASBGSKGBFTYTAVTALKTGVRHLDCAWVYLNEGEVGEIRDFLKENPSVKR 75
Db	RT 2 EDIFVCTKVNWHLRVEDVLSIDSLSKRLGDDYDMFLVHWPINAENGQGEPKIGPDG 135	RT 2 EDIFVCTKVNWHLRVEDVLSIDSLSKRLGDDYDMFLVHWPINAENGQGEPKIGPDG 135
Qy	RT 3 61 EDLFICLKVNWHLRHEEVKRSFENSLSKNGFLDYLFLVHWPINAEXGKGDYQPKIGPDG 120	RT 3 61 EDLFICLKVNWHLRHEEVKRSFENSLSKNGFLDYLFLVHWPINAEXGKGDYQPKIGPDG 120
Db	RT 4 QHGYO PRELIMINARY; PRT; 310 AA.	RT 4 QHGYO PRELIMINARY; PRT; 310 AA.
RA	de Vries R.P.; "Glycerol dehydrogenase, encoded by gldB is essential for osmotolerance in Aspergillus nidulans"; Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.	RA
DR	EMBL; AU196025; CAD42640; 1.	DR
KW	Oxidoreductase.	KW
SQ	SEQUENCE 325 AA; 36981 MW; BFADELCB4DA15F14 CRC64;	SQ
Qy	Query Match 136 RT 1 KVTILDLTENPEPTMAMETYEDRKARSIGVSNTIADLEKMS 195	Query Match 136 RT 1 KVTILDLTENPEPTMAMETYEDRKARSIGVSNTIADLEKMS 195
Db	Best Local Similarity 82.5%; Score 1471; DB 3; Length 325;	Best Local Similarity 82.5%; Score 1471; DB 3; Length 325;
Qy	Matches 268; Conservative 28; Mismatches 29; Indels 0; Gaps 0;	Matches 268; Conservative 28; Mismatches 29; Indels 0; Gaps 0;
Db	RT 2 1 MSGSGTFTLNSGVKIPGVRGTFPASEGSGETYTAVTALKTGVRHLDCAWVYLNEGEVG 60	RT 2 1 MSGSGTFTLNSGVKIPGVRGTFPASEGSGETYTAVTALKTGVRHLDCAWVYLNEGEVG 60
Qy	RT 3 61 EGIRDFLKENPSVKREDIFVCTKVNHLRVEDVLSIDSLSKRLGDDYDMFLVHWPINA 120	RT 3 61 EGIRDFLKENPSVKREDIFVCTKVNHLRVEDVLSIDSLSKRLGDDYDMFLVHWPINA 120
Db	RT 4 61 DAVRDFLKENPSVKREDIFVCTKVNHLRVEDVLSIDSLSKRLGDDYDMFLVHWPINA 120	RT 4 61 DAVRDFLKENPSVKREDIFVCTKVNHLRVEDVLSIDSLSKRLGDDYDMFLVHWPINA 120
Qy	RT 5 121 AEKNGQEPKPGDKYVILKDLTENPEPTWAMEKIYEDRKARSIGVSNTIADLEKMS 180	RT 5 121 AEKNGQEPKPGDKYVILKDLTENPEPTWAMEKIYEDRKARSIGVSNTIADLEKMS 180
Db	RT 6 121 AEKESQDKPKPGDKYVILKDLTEDPKTWOAMEKLYEDKLRARSIGVSNTIADLEKMS 180	RT 6 121 AEKESQDKPKPGDKYVILKDLTEDPKTWOAMEKLYEDKLRARSIGVSNTIADLEKMS 180
Qy	RT 7 181 KFAKYMMPHANOIEHPFLPNEELYOYCFSKINMPVAYSPLGSONQVPTGERSENKLNN 240	RT 7 181 KFAKYMMPHANOIEHPFLPNEELYOYCFSKINMPVAYSPLGSONQVPTGERSENKLNN 240
Db	RT 8 181 KYAKYKPHNQIEHPFLPNEELYOYCFSKINMPVAYSPLGSONQVPTGERSENKLNN 240	RT 8 181 KYAKYKPHNQIEHPFLPNEELYOYCFSKINMPVAYSPLGSONQVPTGERSENKLNN 240
Qy	RT 9 241 EIAEKGGNNTLAQVLIWGLRRGGYVLPKSSNPKRIESNPKSIELSDAEEANAVAKGRH 300	RT 9 241 EIAEKGGNNTLAQVLIWGLRRGGYVLPKSSNPKRIESNPKSIELSDAEEANAVAKGRH 300
Db	RT 10 241 EIAQKGGNNTLAQVLIWGLRRGGYVLPKSSNPARIESNPKSIELSDDEAENAVAKGRH 300	RT 10 241 EIAQKGGNNTLAQVLIWGLRRGGYVLPKSSNPARIESNPKSIELSDDEAENAVAKGRH 300
Qy	RT 11 301 FRFVNMKDTFGDYWPEETAKNLSA 325	RT 11 301 FRFVNMKDTFGDYWPEETAKNLSA 325
Db	SEQUENCE FROM N.A.	RT 11 301 FRFVNMKDTFGDYWPEETAKNLSA 325
RC	STRAIN=ATCC42981;	RC
RA	Iwaki T., Kubota K., Yokose Y., Tamai Y., Watanabe Y.;	RA
RT	"Cloning of glycerol-3-phosphate dehydrogenase genes (ZrgCv1 and ZrgCv2) from the salt-tolerant yeast Zygospacharomyces rouxi.";	RT
RT	ZrgCv2) and glycerol dehydrogenase genes (ZrgCv1 and ZrgCv2) from the salt-tolerant yeast Zygospacharomyces rouxi.";	RT

Result No.	Score	Query	Match	Length	DB ID	Description
1	1514	Q8dyk4 streptococc	Q877a2 aspergillus	30.5	280	16 Q8DYK4
2	1471	Q8sjv1 arabidopsis	Q7z8l1 emericella	30.5	309	10 Q8SJV1
3	1471	Q8aw93 digitalis p	Q74646 gibberella	30.4	315	10 Q8AW93
4	555	Q8gf8 drosophila	Q8gyo2 zygoacchar	30.4	350	5 Q8GF8
5	553	Q81xd1 bacillus an	P74308 synechocyst	30.3	279	16 Q81XD1
6	548	Q81555 bacillus ce	Q42882 schizosach	30.3	279	16 Q81555
7	543	Q8aw92 digitalis p	Q7za52 magnaporthe	30.3	315	10 Q8AW92
8	533	Q8ffw2 medicago sa	Q81bg6 arabidopsis	30.2	316	10 Q8FFW2
9	533	Q88sl1 lactobacill	Q96uh3 coelidiode	30.2	311	10 Q88SL1
10	531	Q818y9 bacillus ce	Q9vt91 drosophila	30.2	311	10 Q818Y9
11	530	Q8944 arabidopsis	Q81b04 bacillus ce	30.2	276	5 Q8944
12	530	Q87pw2 streptococc	Q81mx4 trypanosoma	30.1	315	10 Q87PW2
13	529	Q81md0 bacillus an	Q8m09 arabidopsis	30.1	320	10 Q81MD0
14	529	Q8ld92 arabidopsis	Q8dtr6 streptococc	30.0	280	16 Q8LD92
15	529	Q8est2 oceanobacil	Q8im09 streptococc	30.0	275	16 Q8EST2
16	528	Q9gv41 trypanosoma	Q8jmv2 streptococc	30.2	316	10 Q9GV41
17	528	Q818y9 bacillus ce	Q818y9 trypanosoma	30.2	316	10 Q818Y9
18	527.5	Q818y9 drosophila	Q818y9 streptococc	30.2	316	10 Q818Y9
19	526	Q818y9 bacillus ce	Q818y9 streptococc	30.2	316	10 Q818Y9
20	525.5	Q818y9 drosophila	Q818y9 streptococc	30.2	316	10 Q818Y9
21	525	Q818y9 bacillus ce	Q818y9 streptococc	30.2	316	10 Q818Y9
22	525	Q818y9 drosophila	Q818y9 streptococc	30.2	316	10 Q818Y9
23	525	Q818y9 bacillus ce	Q818y9 streptococc	30.2	316	10 Q818Y9
24	525	Q818y9 drosophila	Q818y9 streptococc	30.2	316	10 Q818Y9
25	523	Q818y9 bacillus ce	Q818y9 streptococc	30.2	316	10 Q818Y9
26	522.5	Q818y9 drosophila	Q818y9 streptococc	30.2	316	10 Q818Y9
27	522.5	Q818y9 bacillus ce	Q818y9 streptococc	30.2	316	10 Q818Y9
28	522	Q818y9 drosophila	Q818y9 streptococc	30.2	316	10 Q818Y9
29	521	Q818y9 bacillus ce	Q818y9 streptococc	30.1	320	10 Q818Y9
30	519	Q818y9 drosophila	Q818y9 streptococc	30.0	320	10 Q818Y9
31	518.5	Q818y9 bacillus ce	Q818y9 streptococc	30.0	275	16 Q818Y9
32	518.5	Q818y9 drosophila	Q818y9 streptococc	30.0	309	10 Q818Y9
33	518.5	Q818y9 bacillus ce	Q818y9 streptococc	29.9	274	16 Q818Y9
34	516.5	Q818y9 drosophila	Q818y9 streptococc	29.8	320	5 Q9VZ8
35	516	Q818y9 bacillus ce	Q818y9 streptococc	29.8	279	16 Q818Y9
36	516	Q818y9 drosophila	Q818y9 streptococc	29.8	280	16 Q818Y9
37	516	Q818y9 bacillus ce	Q818y9 streptococc	29.8	280	16 Q818Y9
38	515.5	Q818y9 drosophila	Q818y9 streptococc	29.8	304	10 Q818Y9
39	508.5	Q818y9 bacillus ce	Q818y9 streptococc	29.4	281	16 Q818Y9
40	507.5	Q818y9 drosophila	Q818y9 streptococc	29.3	276	5 Q8I15
41	506.5	Q818y9 bacillus ce	Q818y9 streptococc	29.3	274	5 Q965C7
42	503	Q818y9 drosophila	Q818y9 streptococc	29.1	276	16 Q32210
43	503	Q818y9 bacillus ce	Q818y9 streptococc	29.1	310	10 Q9FTN7
44	500	Q818y9 drosophila	Q818y9 streptococc	28.9	280	16 Q34678
45	498	Q818y9 bacillus ce	Q818y9 streptococc	28.8	274	16 Q836H3

ALIGNMENTS

RESULT 1
Q877a2 PRELIMINARY; PRT; 325 AA.
ID Q877a2/ AC Q877a2/
DT 01-JUN-2003 (TREMBrel. 24, Created)
DT 01-JUN-2003 (TREMBrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)
DE Reductase-like Protein.
OS Aspergillus oryzae.
OC Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus
OC NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R:FB40;
RA Akao T., Akeno T., Goto K., Akita O.;
RA "cloning and nucleotide sequencing of cDNA that encodes a reductase-"
RA like protein from Aspergillus oryzae";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; Aldo/ket red; 1.
DR PRINTS; PR00059; Aldo/ket red; 1.
DR PRODOM; PD000388; Aldo/ket red; 1.
DR PROSITE; PS00778; ALDOKSTO REDUCTASE 1; 1.
SQ SEQUENCE D5ED3898DOC4BD9 CRC64;

Query Match Score 1514; DB 3; Length 325;
Best Local Similarity 85.8%; Score 1514; DB 3; Length 325;
Matches 22; Mismatches 24; Indels 0; Gaps 0;

1 MSNGKTFITLSNGVKIPGVGFGTFASEGSKGETYTAVTALKTGYRHLDCAWYLYNEGBVG 60
1 EGIRDFLKPKPSVREDIFVCTKVNHLHRVEDYLWSDDSLRKLGLXYYDMLLVHWTIA 120
1 DGHDFLKKNPSPVREDIFVCTKVNHLHRPEDQYSVSDNSLKRLLDXYDVLFLVHWPIA 120

Qy 1 MSNGKTFITLSNGVKIPGVGFGTFASEGSKGETYTAVTALKTGYRHLDCAWYLYNEGBVG 60
Db 1 MSNGKTFITLSNGVKIPGVGFGTFASEGSKGETYTAVTALKTGYRHLDCAWYLYNEGBVG 60

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1514	Q877a2 aspergillus	Q877a2 aspergillus	325	3 Q877a2	
2	1471	Q878l1 emericella	Q7z8l1 emericella	325	3 Q7z8l1	
3	1471	Q74646 gibberella	Q74646 gibberella	325	3 Q74646	
4	555	Q8gyo2 zygoacchar	Q8gyo2 zygoacchar	321	3 Q8gyo2	
5	553	P74308 synechocyst	P74308 synechocyst	31.9	16 P74308	
6	548.5	Q42882 schizosach	Q42882 schizosach	31.7	3 Q42882	
7	543	Q7za52 magnaporthe	Q7za52 magnaporthe	31.4	3 Q7za52	
8	533	Q81bg6 arabidopsis	Q81bg6 arabidopsis	30.8	3 Q81bg6	
9	533	Q96uh3 coelidiode	Q96uh3 coelidiode	31.0	3 Q96uh3	
10	531.5	Q9vt91 drosophila	Q9vt91 drosophila	30.7	5 Q9vt91	
11	530.5	Q81b04 bacillus ce	Q81b04 bacillus ce	30.6	279 Q81b04	
12	530	Q9hgx9 zygosacchar	Q9hgx9 zygosacchar	31.0	3 Q9hgx9	
13	530	Q8gxw0 arabidopsis	Q8gxw0 arabidopsis	30.6	322 Q8gxw0	
14	529.5	Q9sv72 arabidopsis	Q9sv72 arabidopsis	30.6	309 Q9sv72	
15	529	Q80945 arabidopsis	Q80945 arabidopsis	30.6	290 Q80945	
16	528	Q8E460 streptococc	Q8E460 streptococc	30.5	280 Q8E460	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

XYL1_PACTA STANDARD; PRT; 318 AA.
 AC P76136;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pachysolen tannophilus (yeast) H-dependent xylose reductase (EC 1.1.1.-) (XR).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetidae; Saccharomycetaceae; Pachysolen.
 NCBI_TaxID=4918;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y-2460 / ATCC 32691;
 RX MEDLINE=970825742; PubMed=8923742;
 RA Bolen P.L., Hayman G., Shepherd H.S.;
 RT "Sequence and analysis of an aldose (xylose) reductase gene from the
 xylose-fermenting yeast Pachysolen tannophilus.";
 RL yeast 12:1367-1375(1996).
 RN [2]
 RP SEQUENCE OF 1-62.
 RA Bolen P.L., Bietz J.A., Detroyn R.W.;
 RT "Aldose reductase in the yeast Pachysolen tannophilus: purification,
 characterization and N-terminal sequence";
 PT Biochemical Bioeng. Symp. 15:129-148(1985).
 RL FUNCTION: Reduces xylose into xyitol.
 CC -!- PATHWAY: D-xylose degradation.
 CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL curatorium on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: U00706; AAC:9526.1; -.
 DR HSSP: P14550; ZALR.
 DR InterPro: IPI0001395; Aldo/ket_red.
 DR PFam: PF00248; aldo_ket_red_1.
 DR PRNTS; PR00069; ALDOKETORED.
 DR PROBOM: PD000288; Aldo/ket_red.
 DR PROBOM: PD000289; Aldo/ket_red.
 DR PROSITE; PS00798; ALDOKETO_REDUCTASE_1;
 DR PROSITE; PS00062; ALDOKETO_REDUC_2;
 DR PROSITE; PS00063; ALDOKETO_REDUC_3; FALSE_NEG.
 KW Xylose oxidoreductase; Xylose metabolism; NAD.
 FT ACT SITE 48 HYDROGEN_BOND_DONOR (BY SIMILARITY).
 SQ SEQUENCE 318 AA;
 36122 MW; 891FC6746LD231EC CRC64;
 Query Match 28.1%; Score 466; DB 1; Length 318;
 Best Local Similarity 38.2%; Pred. No. 6.2e-31;
 Matches 125; Conservative 51; Mismatches 117; Indels 34; Gaps 11;
 Qy 7 FTLLANGVK1PGVGRTGTPASEGSKGETYAVTTALKTYRHLDCAWTYLINEGYEGEGRDF 66
 Db 6 YTLNNGRK1PAIGMGCKWLENAADMVYAAI----KEGYRLFOCADYGNEKEVGEGLNRA 61
 Qy 67 LKENPSVKRBDIFVCTKWNHHLARYDVLSIDSRLRLGQYVDMFLVHPIAAEKNGQ 126
 Db 62 IKDG-LVTKDLDLFTSKLNWNFHAKENYKKALMKSLSDFNLNDYFDLXLMHPLSF-KFVP 119
 Qy 127 GEPKIGP----DSKXYTILKDLTENPEP--TWRAMEKXYEDRKARTSGVSNWTALEK 178
 Db 120 FEEXKPPGFYWCDDKPIYEDV----PIETTRAMENLVDELVKSIGNSVSGILED 174
 Qy 179 MSKPKAYKVMHANOQIBHPFLPNEEBLVOYCFKNIMPVAYSPGS-----QNQVPTGER 232
 Db 175 LIKARIKIKASLQ.EHHPYLQCNLVEAQLKGIVVTGYSNGLPSFILEGNETAKTQP 234
 Qy 233 VSENKTINELAAEKGGNTLAQVLIAGWLBRGGYYVLKPSSNPKRIESNF--KSEIELSADFE 290
 Db 235 LYENRTITTAAKHGKTPFQVLLRWNRORGIAIPKSTPNTLAVNLHVDPLTLKEDFE 294

Proc. Natl. Acad. Sci. U.S.A. 99:16899–16903 (2002).

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

DECEMBER 15

DR	EMBL; U13692; AAB60687.1; JOINED.	SEQUENCE FROM N.A. STRAIN=ICR X Swiss Webster; TISSUE=Liver;
DR	EMBL; U13693; AAB60687.1; JOINED.	RC Iwata T., Carpenter D.J., Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; M32818; AAA31160.1; -.	RA Iwata T., Carpenter D.J., Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; U12316; AAA50533.1; -.	RL Daoudal S., Berger M., Pailhous E., Tournaire C., Veysiere G., Jean C.; Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; J05048; AAA31157.1; -.	[3]
DR	PRINTER; A34406; A34406.	[4]
DR	HSSP; P15121; 2ACQ.	SEQUENCE FROM N.A. STRAIN=CD-1; TISSUE=Kidney;
DR	InterPro; IPR001395; Aldo/ket red.	RC Medline=98153248; PubMed=9485485;
DR	DR	RA McGowan M.H., Iwata T., Carpenter D.A.; "Characterization of the mouse aldose reductase gene and promoter in a lens epithelial cell line." Mol. Vision 4:2(1998).
DR	Oxidoreductase; NADP; Acetyltransferase; NADP; Acetylation.	RA Louie H.W.Y., Lee F.K., Chung S.S.M., Chung S.K.; "Comparisons of genomic structures and chromosomal locations of the mouse aldose reductase and aldose reductase-like genes." Eur. J. Biochem. 259:726-730(1999).
FT	INIT MET 0	RT RL RN
NP	NP BIND 9	RT RN
FT	ACT SITE 48	RT RN
MOD RES 1	35632 MW; 64D53E6AC0833FFB CRC64;	RT RN
SQ	315 AA;	RT RN
Qy	9 LSNQVXIPGVFGTFASEGSKGSKGETTYAVTALKGYRHLDCAWYLYNNEGVGGIRDPLK 68	SEQUENCE FROM N.A. STRAIN=BALB/C;
6 LYNGAAMPILGLGTWS - RPPGQVTEAKTAIDGYRHIDCATYQNEENVGFLAQEKX 63	RC Medline=98160426; PubMed=10019784;	
Qy	69 ENPSVRSDEFIFVCTWKWNHFLHRYELVWLSDDSSRKLGLDYYDVDFVWPIAAEKGNGGE 128	RA Li H., Nobukuni Y., Gui T., Yabe-Nishimura C.; "Characterization of genomic regions directing the cell-specific expression of the mouse aldose reductase gene." Biochem. Biophys. Res. Commun. 255:759-764 (1999).
Db	64 EQ-VVREELFIVSKWLTSHDSLVKGAAQKINDLKDYLILWPTGFKHGSEYF 122	RT RN
Qy	129 PKIGPDGKTYLKLDTENPEP-----TWRAMEKTYEDRKARSIGVSNTWTADELMK-- 179	RT RN
Db	123 P-----LDAAGNVPFSDTDFLDWBAMBGLVDLGVLKSGVSNFHIIQIERILN 171	RT RN
Qy	180 SKPAXYMPHANOIEIETHPLPNEELVQYCFSKNMVPAYSPLGSONQVPTTGERVS--ENK 237	RT RN
Db	172 KPGLKXYPANQIECHPYLQKLUQYCHQSKIVYTASPLGSDRPAKPEDSLLDP 231	RA Klausner R.D., Collins F.S., Wagner L., Sheinman C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsiieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Casavant T.L., Scheetz T.E., Stapleton M., Soares M.B., Bonaldo M.F., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loquellat N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McErwan P.J., McKeran K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villaon D.K., Muzny D.M., Soergren E.J.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Rodriguez A.C., Grinwood J., Green E.D., Dickson M.C., Butterfield Y.S.N., Krzyzwiak M.I., Shalska U., Smalius D.E., Schein J.B., Jones S.J.M., Marra M.A., Marra M.A., Schnierch A., Schein J.B., Jones S.J.M., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
Qy	238 TLNEIAEKGENTLAQYLIANGLRROXYVLPKSSPFRLESNFK--SIEUSDAFEANAV 295	RT RN
Db	232 RIKAIADKHKTTAQVLIRPFMQRNLVWPKSVTPARIAENQVFDPELSSDMTL--L 289	RA Butterfield Y.S.N., Touchman J.W., Schmutz J., Myers R.M., Rodriguez A.C., Grinwood J., Green E.D., Dickson M.C., Schein J.B., Jones S.J.M., Marra M.A., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
Qy	296 ARGRHPR 302	CC -1- CATALYTIC ACTIVITY: Alditol + NAD(P)(+)= aldoze + NAD(P)H.
Db	290 SYNWR 296	CC -1- SUBCELLULAR LOCATION: Cycloplasmic.
RESULT 12	CC -1- TISSUE SPECIFICITY: Abundant in the testis, skeletal muscle and kidney.	
ID _ALDR_MOUSE	STANDARD; PRT; 315 AA.	CC -1- SIMILARITY: Belongs to the aldo/keto reductase family.
AC P45276; OT0130; Q99KC9.	DT 01-NOV-1995 (Rel. 32, Created) 01-FEB-1996 (Rel. 33, Last sequence update) 10-OCT-2003 (Rel. 42, Last annotation update)	CC -1- CATALYTIC ACTIVITY: Alditol + NAD(P)(+)= aldoze + NAD(P)H.
DE Aldose reductase (EC 1.1.1.21) (AR) (Aldehyde reductase).	DE RN	CC -1- SUBUNIT: Morome.
OS ARK1B1 OR ALDR1 OR ALDOR1.	OS Mus musculus (Mouse). OC Eukaryota; Metzoa; Chordata; Craniata; Vertebrata; Buteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI_TaxID=10030;	CC -1- TISSUE SPECIFICITY: Abundant in the testis, skeletal muscle and kidney.
RP STRAIN=C57BL/6; TISSUE=Kidney.	RP Medline=95154335; PubMed=7851421; RA Gu T., Tanimoto T., Kokai Y., Nishimura C.; "Presence of a closely related subgroup in the aldo-ketoreductase family of the mouse." Eur. J. Biochem. 227:448-453 (1995).	CC -1- This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no

KW	Oxidoreductase; NADP; Acetylation; 3D-structure.	6 LYTGAAKMPILGLGTWS- - PPGKVTEAVKVAIDIGYRHIDCAHYQNENEVGLGLOQEKU-
FT INIT MET	0	18 NADP (POTENTIAL).
FT NP_BIND	9	48 HYDROGEN BOND DONOR (BY SIMILARITY).
FT ACT_SITE	48	1 HYDROGEN BOND ACCEPTOR (BY SIMILARITY).
FT MOD_RES	1	1 ACETYLATION.
FT DISULFID	298	303 D -> N (IN REF. 2).
FT CONFLICT	98	98
FT STRAND	3	5
FT TURN	7	8
FT STRAND	11	13
FT STRAND	15	15
FT STRAND	17	18
FT TURN	20	21
FT HELIX	24	37
FT STRAND	38	38
FT STRAND	41	43
FT HELIX	46	48
FT HELIX	51	63
FT TURN	64	65
FT HELIX	69	71
FT STRAND	73	78
FT HELIX	80	82
FT HELIX	85	87
FT HELIX	88	99
FT TURN	100	100
FT STRAND	104	104
FT STRAND	106	109
FT STRAND	115	115
FT STRAND	115	124
FT TURN	124	124
FT STRAND	126	126
FT STRAND	130	130
FT STRAND	132	132
FT HELIX	137	150
FT TURN	151	151
FT STRAND	153	153
FT STRAND	156	159
FT HELIX	163	170
FT TURN	173	174
FT TURN	173	174
FT STRAND	181	185
FT STRAND	187	187
FT TURN	188	189
FT STRAND	190	190
FT HELIX	193	202
FT TURN	203	203
FT STRAND	205	209
FT TURN	211	212
FT TURN	215	216
FT TURN	218	219
FT TURN	222	223
FT HELIX	227	229
FT HELIX	231	240
FT TURN	241	241
FT HELIX	244	254
FT TURN	255	256
FT STRAND	258	259
FT STRAND	262	262
FT HELIX	266	271
FT TURN	272	277
FT HELIX	282	289
FT TURN	290	291
FT HELIX	301	303
FT TURN	304	305
FT TURN	307	308
FT HELIX	310	312
SQ SEQUENCE	315 AA;	3573 MW; F3E18FB14E225C92 CRC64;
Query Match	28.2% Score 488.5; DB 1; Length 315;	9 LSNGYTKI PGYCFGTASESEGSKGERTAATALKGYRHDCAVNTLNLSGEVGIGIRDLK 68
Best Local Similarity	36.9%; Pred. No. 3 9e-31;	Matches 116; Conservative 61; Mismatches 122; Indels 15; Gaps 8;
Matches		
Db	Qy	6 LYTGAAKMPILGLGTWS- - PPGKVTEAVKVAIDIGYRHIDCAHYQNENEVGLGLOQEKU-
Db	Qy	69 ENPSVREDFIVCTKVNHLHRYDVLWSDDSLKRLGDDYVDMFLVHVKPIAAKXNGQE 128
Db	Qy	63 QGQQVREDFLPIVSQKWCOTDHEKVLKGACOTLRLDKLQDLYLHWPTGFK--PSK 119
Db	Qy	129 PKIGPDKYVILKDLTENPEPTWANEKIYEDKARSIGVSNTTADLERN--SKEAKYM 186
Db	Qy	120 DPFLPLGDGNVVPDSEDFVE-TWEAMEELVDEGL-YKAIGVSNEFHLOVEKILNKPGLYK 178
Db	Qy	187 PHANQIEITHPPFLPNEELVQYCFSKNTIMPAYSPJGSQNQVPTTGERVS--ENKTLINEIAE 244
Db	Qy	179 PAVNQEVHPVLTQEMLIECKSKGIVVTTAXSPLOSSPDRPWAKPDPSLLEDPRKAIKA 238
Db	Qy	245 KGGNTLAQVLLAWGLRGYVLPKSSNPKRIESNPK--STELSDADEFAINAVAKGRHFR 302
Db	Qy	239 KYNKTTAQVLLRFPMQRNLIVPKSVTPERIAEFQVFDFELSPPDMNTL--LSINRNWR 296
Db	Qy	303 FVNMKUTFGYDVWP 316
Db	Qy	297 VCALMSCAHXRDP 310
RESULT 11		
ALDR_RABBIT	ID ALDR_RABBIT	STANDARD;
AC	AC P15122;	PRT;
DT	DT 01-APR-1990 (Rel. 14, Created)	
DT	DT 01-OCT-1995 (Rel. 34, Last sequence update)	
DT	DT 28-FEB-2003 (Rel. 41, Last annotation update)	
DE	DE Aldose reductase (EC 1.1.1.21) (AR) (Aldehyde reductase).	
GN	GN AKR1B1.	
OS	OS Oryctolagus cuniculus (Rabbit).	
OC	OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagidae.	
NCBI_TaxID	NCBI_TaxID=9986;	
RN	RN [1]	
RP	RP SEQUENCE FROM N.A.	
RC	RC TISSUE=Spleen;	
RX	RX MEDLINE=95024191; PubMed=7938022;	
RX	RX MEDLINE=89380313; PubMed=2506183;	
RA	RA Garcia-Perez A., Martin B., Murphy H.R., Uchida S., Murer H., Ferraris J.D., Williams C.K., Martin B.M., Burg M.B., Burg R.M.B., Cowley B.D. Jr., Handler J.S., Burg M.B.; "Cloning, genomic organization, and osmotic response of the aldose reductase gene,"	
RA	RA Proc. Natl. Acad. Sci. U.S.A. 91:10742-10746 (1994).	
RL	RL RN [2]	
RP	RP SEQUENCE OF 14-315 FROM N.A.	
RC	RC TISSUE=Kidney;	
RX	RX MEDLINE=89380313; PubMed=2506183;	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	
CC	CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.	
CC	CC -!- CATALYTIC ACTIVITY: Alditol + NAD(B) (+) = aldose + NAD(P)H.	
CC	CC -!- SUBUNIT: Monomer.	
CC	CC -!- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	CC -!- SIMILARITY: Belongs to the aldol/keto reductase family.	

PRINTS; PRO0069; ALDKETREDTASE.
DR PROD0; PD000298; Aldo/*ket*_red.; 1.
DR PROSITE; PS001798; ALDOKETO_REDUCTASE_1;
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_2;
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; 1.
KW Oxidoreductase; NADP.
ACT SITE 48 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
SQ SEQUENCE 310 AA; DECCAD94BF556D9F CRC64;

Query Match 29.0%; Score 501.5; DB 1; Length 310;
Best Local Similarity 34.8%; Pred. No. 3.7e-32; Gaps 8;
Matches 110; Conservative 73; Mismatches 118; Indels 15; Gaps 8;

Qy 6 TFTLNSGVKIPGVGFTRASEGSKGETYPAVTPALKTGVRHLDCAWYLYNEGEVGEGIRD 65
Db 3 TVVLLSGVMEPVIGLMLAE - KDELKVILNARIKGTHFDAAHYKEADVGELAE 60
Qy 66 FLKENPSPVREDFIFCTKWNHLHYEDVLSIDDSKLRLGDDYDVMFLVHMPTKHNA 125
Db 61 AFKTS-LVKRBEELFTTKWNSDHG--VTEACKNSLELQIDBLDVLYHMPTKHNA 117
Qy 126 QGEPK- -IGDGKXYLKLDTENPEPTKAYEKEIYEDKARSTGVSNTTADLEMSKEA 183
Db 118 IGKTAISLGD--KVLDDIVTIS-QQTWSMKTIVSLGIVRSGLSNVLFUTRDCLAYS 175
Qy 184 KMMPHANOQIEHHPFLPNELVQYCEFSKNIMPVAVSPIGSQNYQPTTGERVS - ENKTLINE 241
Db 176 KIKPAVSPOTPHYQFDSLVKFCMKHGVLPTATPLGGAAANAKDMFGSSVSPDDPVND 235
Qy 242 IAEKGNTLAQVLIANGLRRGYVVLPKSSNPKRLESNFKSIS -- LSDADFEATNAVAKGR 299
Db 236 VAKYGKSVAQICLWRGICRKTAVTPKSQKIQRLKENLEVLEFQLSDSDMQLIYSI - DR 293
Qy 300 HFRFNMKQDGTGVDW 315
Db 294 KYRTSLPSKTRWGLDYY 309

RESULT 8

ALDR RAT STANDARD; PRT; 315 AA.

ID P0743; 1888 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aldose reductase (EC 1.1.1.21) (AR) (Aldehyde reductase).
GN ARB1 OR ALDR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBIX_TaxID:10116;

RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN=Sprague-Dawley; TISSUE=Lens;
RX MEDLINE=8722765556; PubMed=311886;
RA Carper C., Nishimura C., Shinohara T., Dietzchold B., Wistow G.,
RA Craft C., Kador P., Kinoshita J.H.;
RT "Aldose reductase and p-cresylallin belong to the same protein
superfamily as aldehyde reductase.",
PESS Lett. 220:209-213 (1987).
RN SEQUENCE FROM N.A.
RX MEDLINE=92084116; PubMed=1748296;
RA Graham C.E., Szpirer C., Levan G., Carper D.;
RT "Characterization of the aldose reductase-encoding gene family in
rat." Gene 107:259-267 (1991).
RN [3] SEQUENCE OF 155-168 AND 204-209.
RT TISSUE=Astrocyte;
RX MEDLINE=96007949; PubMed=7498172;
RA Laeng P., Bouillon P., Taupenot L., Labourette G.;
RT "Long-term induction of an aldose reductase protein by basic

RT fibroblast growth factor in rat astrocytes in vitro."
RL Electrophoresis 16:1240-1250 (1995).
CC -!- FUNCTION: Catalyzes the NADPH-dependent reduction of a wide variety of carbonyl-containing compounds to their corresponding alcohols with a broad range of catalytic efficiencies.
CC -!- CATALYTIC ACTIVITY: Alditol + NAD(P)(+)= aldose + NAD(P)H.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the aldo/keto reductase family.

RT This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC DR X05884; OAA29308.1; -.
DR EMBL; MG0322; AAA40721.1; -.
DR PIR; A60603; A60603.
DR HSSP; P1521; 2ACO.
DR InterPro; IPR01395; Aldo/ket_red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR PRINTS; PRO0069; ALDKETERETASE.
DR PRODOM; PDO288; Aldo/ket_red; 1.
DR PROSITE; PS00062; ALDOKETO_REDUCТАSE_2; 1.
DR PROSITE; PS00033; ALDOKETO_REDUCТАSE_3; 1.
DR PROSITE; PS00798; ALDOKETO_REDUCТАSE_1; 1.
DR KW Oxidoreductase; NADP; Acetylation.
FT INIT MET 0 0 BY SIMILARITY.
FT NP BIND 9 18 NADP (POTENTIAL)
FT ACT-SITE 48 48 HYDROGEN-BOND DONOR (BY SIMILARITY).
FT MOD RES 1 1 ACTYLATION (BY SIMILARITY).
SQ SEQUENCE 315 AA; 356656 MW; 572941A154BC1202 CRC64;

Query Match 28.8%; Score 499.5; DB 1; Length 315;
Best Local Similarity 39.4%; Pred. No. 5.4e-32;
Matches 121; Conservative 51; N mismatches 106; Indels 29; Gaps 10;

Qy 9 LSNGVLPGVOFGTFFSEGSKGETYAVTALKTGRLDAWYLYNEGEVGEGIRDFLK 68
Db 6 LNNGTQPTPLGLTWWK - -PPGQVTEAVKVAIDMGYRHICAQVQNKEVGVALQEKLK 63
Qy 69 ENPVSVKREDIPVCTKWNHLHYEDVLSIDDSKLRLGDDYDVMPLVHMPTAEEKGQGE 128
Db 64 EQ-VVQRQDLFIVSKWCMTFIDQSMWGRAGCQTLSDQLQYDLYLHWPTGF - - - - 115
Qy 129 PKIGPDCKVYVILKDITDENPER - - - - -TWRAMEKTYEDKARSTGVSNTTADLEKMR - - - - 179
Db 116 -KPGPD-YFPF-DAGGNVTSDDTDFTDVTNTAMEQVDEGLVKAIGVSNNNPLQIERILN 171
Qy 180 SKPARYMPHANOQIEHPLNEMELVQYCEFSKANIMPVAKSPLGSQNOVPTTGERVS - ENK 237
Db 172 KEGLKVKPAPVNOQIECPYLQCEKLYCHCKGIVVYAYSPLGSPRWPWAKPEDPSLLEDP 231
Qy 238 TUNIEAERGGNTLQAQVLIANLRLRGVVLPKSSNPRIESENFK - SIELSDADFAINAV 295
Db 232 RIKEIIAKYKNTTAQVIRPFIQRLVVIWKSUTPARIABNEYKDFEESNEDMTI - I 289
Qy 296 AKGRHFR 302
Db 290 SYNNRNRR 296

RESULT 9

P100_BIMA STANDARD; PRT; 284 AA.
ID P100_BIMA
AC P22045;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable reductase (EC 1.1.1.-).

RESULT 6

ID	ALDI_MOUSE	STANDARD;	PRT;	315 AA.
AC	P21370;			
DT	01-MAY-1991 (Rel. 1.8, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DT	Aldose reductase protein 1 (BC 1.1.1.21) (AR) (Aldo-keto reductase) (VAS deferens androgen-dependent protein) (MyD) (ARR17 OR AVDP)			
GN	Mus musculus (Mous).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.			
OX	NCB_TaxID=10090;			
RP	SEQUENCE FROM N.A.:			
RC	TISSUE=vas deferens;			
RX	Medline:92345100; PubMed:1637719;			
RA	Pailhoux E.A., Veissiere G.M., Fabre S., Tournaire C., Jean C.G.;			
RT	"The genomic organization and DNA sequence of the mouse vas deferens androgen-regulated protein gene.";			
RL	J. Steroid Biochem. Mol. Biol. 42:561-568(1992).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RX	STRAIN=CD-1; TISSUE=vas deferens;			
RA	Medline:91060614; PubMed=2123194;			
RT	"Androgen-dependent protein from mouse vas deferens cDNA cloning and protein homology with the aldo-keto reductase superfamily.";			
RT	J. Biol. Chem. 265:19932-19936 (1990).			
RL	-!- FUNCTION: The role of MyD in sperm maturation and storage may be related to its potential capacity to produce fructose, or MyD may play an osmoregulatory role by producing sorbitol.			
CC	DE (Aldose-6-phosphate reductase [NADPH]) (NADP-S6PDH).			
CC	S6PDH.			
CC	OS Malus domestica (Apple) (Malus sylvestris).			
CC	CN Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Maloideae; Maloidea; Malus.			
CC	OC			
CC	NCBI_TaxID=3750;			
CC	RN [1]			
CC	RN [2]			
CC	RP SEQUENCE FROM N.A.			
CC	RA Kanayama Y., Mori H., Imaeiki H., Yamaki S.;			
CC	RT "Nucleotide sequence of a cDNA encoding NADP-sorbitol-6-phosphate dehydrogenase from apple."			
CC	RT Plant Physiol. 100:1607-1608 (1992).			
CC	RP SEQUENCE FROM N.A.			
CC	RA Bains H.S., Tao R., Uatsu S.L., Dandekar A.M.;			
CC	RT "Genomic nucleotide sequence of NADP sorbitol-6-phosphate dehydrogenase (NADP-S6PDH) gene from apple.";			
CC	RL (In) Plant Gene Register PG98-193.			
CC	-!- FUNCTION: Synthesizes sorbitol-6-phosphate, a key intermediate in the synthesis of sorbitol which is a major photosynthetic product in many members of the Rosaceae family.			
CC	-!- CATALYTIC ACTIVITY: Alditol + NAD(P)H = aldose + NAD(P)H.			
CC	-!- SUBUNIT: Monomer (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Cyttoplasmic (By similarity).			
CC	-!- INDUCTION: Castration resulted in a marked decrease in the level of the mRNA coding for the protein, whereas administration of testosterone to castrated males resulted in a marked increase.			
CC	-!- SIMILARITY: Belongs to the aldo/keto reductase family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; M81448; AAA3974_1.			
DR	EMBL; J05663; AAA3973_1; -.			
DR	PIR; A3790; A3790.			
DR	HSSP; P4537; IFRB.			
DR	MGI; MGI:101948; Akrib7.			
DR	InterPro; IP0001398; Aldo/ket_red.			
DR	PRINTS; PR00048; aldo_ket_red.			
DR	PRODOM; PD000288; Aldo/ket_red.			
DR	PROSITE; PS00062; ALDOKETO REDUCTASE_2; 1.			
DR	PROSITE; PS00063; ALDOKETO_REDUCtASE_2; 1.			
DR	Oxidoreductase; NADP.			
FT	INT_MET 0 0 BY SIMILARITY.			
FT	ACT SITE 48 48 HYDROGEN-BOND DONOR (BY REF. 2).			
FT	CONFLICT 93 93 Q->D (IN REF. 2).			
SQ	SEQUENCE 315 AA; 35857 MN; EA3DS28294AF8542 CRC64;			
CC	DR EMBL; D11080; BAA01853_1.			
CC	DR EMBL; AF057134; AAC97607_1; -.			
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	CC DR PIR; T17013; T17013.			
CC	CC DR HSSP; P14550; ZAIR.			
CC	CC DR InterPro; IPR001395; Aldo_ket_red.			
CC	CC Pfam; PF00248; aldo_ket_red; 1.			
CC	CC DR EMBL; D11080; BAA01853_1.			
CC	CC DR PIR; T17013; T17013.			
CC	CC DR HSSP; P14550; ZAIR.			
CC	CC DR InterPro; IPR001395; Aldo_ket_red.			
CC	CC Pfam; PF00248; aldo_ket_red; 1.			

Query Match 29.1% Score 503; DB 1; Length 315;

Best Local Similarity 38.7% Pred. No. 2.8e-32;

Matches 123; Conservation 58; Mismatches 121; Indels 16; Gaps 9;

QY 6 TFT-LSNGYKIPGVGFGTPASESGSKGETYAVTALKTGYRLUDCAWVYLINEGEVGEIGIR 64

DR	GO; GO:0016491; F:oxidoreductase activity; IDA.	RX	MEDLINE=99117109; PubMed=9920381;
DR	GO; GO:0019566; P:arabinose metabolism; IMP.	RA	Kim S.T., Huh W.K., Lee B.H., Kang S.O.; "D-arabinose dehydrogenase and its gene from <i>Saccharomyces cerevisiae</i> ."
DR	InterPro; IPR001395; Aldo/ket red.	RT	"D-arabinose dehydrogenase."
DR	Pfam; PF00248; aldo/ket red.	RL	DR Biochim. Biophys. Acta 1429:29-39(1998).
DR	PRINTS; PD00060; ALDKETORETASE.	CC	-I- FUNCTION: Catalyzes the oxidation of D-arabinose, L-xylose, L-fucose and L-galactose in the presence of NADP+.
DR	ProDom; PD00058; Aldo/ket red; 1.	CC	-I- CATALYTIC ACTIVITY: D-arabinose + NAD(P) (+) = D-arabinono-1,4-lactone + NAD(P)H.
DR	PROSITE; PS00062; ALDOKETO_REDUCTASE_-2; 1.	CC	-I- SUBUNITS: Heterodimer of a heavy chain and a light chain.
DR	PROSITE; PS00063; ALDOKETO_REDUCTASE_-3; 1.	CC	-I- SUBCELLULAR LOCATION: Cytoplasmic.
DR	PROSITE; PS00798; ALDOKETO_REDUCTASE_1; 1.	CC	-I- MISCELLANEOUS: Exhibits maximum activity at pH 10.0 and around 30 degrees Celsius.
KW	Oxidoreductase.	CC	-I- SIMILARITY: Belongs to the aldd/keto reductase family.
FT	ACT SITE 56 HYDROGEN-BOND DONOR (BY SIMILARITY).	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
SQ	SEQUENCE 312 AA; 34755 MW; 078F92FACEC70E CRC64;	CC	EMBL; Z36018; QAA85107.1; -.
Query Match	Best Local Similarity 32.1%; Score 555; DB 1; Length 312;	CC	EMBL; M9550; AAA35037.1; -.
	Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;	CC	PIR; S46020; S46020.
Qy	2 SNGKTFITLNGVKTPGFGTPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGE 61	CC	HSSP; P06632; 1HW6.
Db	7 NSSATLKLKCASTEVLPFGTWSYDNG-YHYSVIAAKAGYRHIDAAVLYNEEVGR 64	CC	Germline; 138692; -.
Qy	62 GIRDFLKENSVKRDLFIVCTKWNHLHAYEDVWLSIDSLSLRKGIDVDMFLVHWPPLA 121	CC	SGD; S000353; ARA1.
Db	65 AIKD---SGVPREBIFITKLKGTEQR-DPEAALAKSLKRPLGIDVYLMMHWPPL 121	CC	GO; GO:0045290; F:D-arabinose 1-dehydrogenase [NAD(P)] activity; IDA.
Qy	122 EKNGOGEPKIGPGDKXVILKDLTE-----NPEPTWRAMEKUYEDRKARSIGVSNW 171	CC	GO; GO:0010595; F:carbonate metabolism; IDA.
Db	118 KTDRLV----TGDGVLCITITLEQSTDTKETNFNFIKWLWQELPTGKTAVGVSNF 171	CC	InterPro; IPR01395; Aldo/ket red.
Qy	172 TIADELMK--SKFAKYMMPHANOIETHPFLPNEEUVOYCFSTKNMPVAYSPLGSQNQVPT 229	CC	PFam; PF00248; aldo_ket_red; 1.
Db	172 SINKKLELSPNPKVVPATNQIETPLPQDEQDIAFCBKGIIVVEASPPFGAN_APL 230	CC	PRINTS; PRO0069; ALDKETORETASE.
Qy	230 GERVSENKTNLIAFKGGNTLAQVLIAMURGVVILKKSNNKPIESNEKSFELSDDAP 289	CC	Prodrom; PD00288; Aldo/ket red; 1.
Db	231 KE----QALIDMKXKHGJEPQAQTISNSIQRGTVVLAKSVNPERIVSNFKFTLPEPDF 285	CC	DR Prosite; PS00738; ALDOKETO_REDUCTASE_1; 1.
Qy	290 EAINAVAKGHRH-RPYNMK 307	CC	DR Prosite; PS00062; ALDOKETO_REDUCTASE_2; 1.
Db	286 KTISNLISKVHGTKRYDMDK 304	CC	DR Prosite; PS00062; ALDOKETO_REDUCTASE_3; 1.
		CC	KW Oxidoreductase; NAD.
		FT	HYDROGEN-BOND DONOR (BY SIMILARITY).
		FT	ACT SITE 71; 71
		FT	AIAKAGY -> LQSKDNN (IN REF. 2).
		FT	CONFLICT 56 62
		FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
		FT	FT (IN REF. 2).
		FT	7 FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
		FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RESULT 5	ARA1 YEAST STANDARD; PRT; 344 AA.	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
ID	ARA1 YEAST	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
AC	P38115	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
DT	01-OCT-1984 (Rel. 30, Created)	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
DT	01-OCT-1994 (Rel. 30, Last sequence update)	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
DT	15-NAR-2004 (Rel. 43, Last annotation update)	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
DE	D-arabinose dehydrogenase [NAD(P)+] heavy chain (EC 1.1.1.117).	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
GN	ARA1 OR YBR149W OR YBR1127.	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
OS	Eukaryotes cerevisiae (Baker's yeast).	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
OC	Ascomycota; Saccharomycotina; Saccharomyces;	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
NCB_TaxID	NCB_TaxID=4932;	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RN	[1]	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RP	SEQUENCE FROM N.A.	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RC	STRAN=2:88C;	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RA	Entian K.-D., Koettner P., Rose M., Becker J., Grey M., Li Z., Niegemann E., Scheik-Groeninger R., Servos J., Wehner E., Wolter R., Brendel M., Buer J., Braun H., Dern K., Duesterhus S., Grunbein R., Hedges D., Kiesau P., Korol S., Krems B., Proft M., Siegers K., Baur A., Boles E., Miosga T., Schaff-Gerstenschlaeger I., Zimmermann F. K.; Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RN	[2]	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RP	SEQUENCE FROM N.A.	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RX	MEDLINE=91171289; PubMed=2005616;	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RA	Martinez-Soriano J.-P., Wong W.M., van Ryk D.I., Nazar R.N.; "A widely distributed 'CAT' family of repetitive DNA sequences."	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RN	[3]	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66
RP	CHARACTERIZATION, AND SEQUENCE OF 7-20.	FT	FTLSNGVKIPGEGVGFPTASEGSKGETTYATTALKTGTRHLDCAWYLNNEGEVGF 66

DR	EMBL; X90518;	CA462107.1;	-;
DR	SGD; S0015616;	Gcv1	-;
DR	GO; GO:0004033;	Aldo-keto reductase activity; IDA.	
DR	GO; GO:0096511;	P: salinity response; IEP.	
DR	InterPro; IPR01305;	Aldo/ket red.	
DR	Pfam; PF00248;	aldo/ket red_1.	
DR	PRINTS; PR00069;	ALDOKETREDTASE.	
DR	Prodrom; PD00288;	Aldo/ket red_1.	
DR	PS00062;	ALDOKETO REDUCTASE_2;	1.
DR	PROSITE; PS00063;	ALDOKETO REDUCTASE_3;	1.
DR	PROSITE; PS00798;	ALDOKETO_REDUCtASE_-1.	
RW	Oxidoreductase.		
FT ACT SITE	56	HYDROGEN BOND DONOR (BY SIMILARITY).	
SQ SEQUENCE	312 AA;	567939 MW; 779BD31C228334A4 CRC64;	
Query Match Best Local Similarity	40.8%;	Score 574.5; DB 1;	Length 312;
Matches 127;	Conservative 54;	Pred. No. 6.8e-38;	
	Mismatches 85;	Indels 45;	Gaps 7;
Qy	5 KTFTLNSGKVIPKGYGFCTFASEGSKGKETYTAATPTALKTGXRHLDCAWYLYNEGEVGEGRIR	64	
Db	10 KILSLNTAQIOPQIGLCTWQSKE-ENDAKAVITALDKXRHIDTAATYRNEDDQGQAIIK	67	
Qy	65 DFLKENPSVKREDIFVCFTKWVNHLLRYEDVYLWSDDSKLRGLDYYDMFLVHPPIAAEKN	124	
Db	68 D-----SGVYRREELFVTKLWCTQHHEPV--ALDQSLKRGLDYYDLYMHNPFLR---	117	
Qy	125 GQGPBKPGDKYVILKDLINEP-----PTNAMEK1YEDRKARSICGV	168	
Db	118 -----DPAYIKNEDLSVPTKDKDSRAVDITNNFKTWTMELQELPKTKTAKVG	168	
Qy	169 SNTIADLEKM-SKFEAKYMPHANQIEHTFLPNELVYCFSKSNIMPAYSPLGSONQV	226	
Db	169 SNTFSINNLKDLLASQGNKLTPAANQEVHPLPDELLNFCKSKGIVIBAYSPLG-----	224	
Qy	227 PTGGERVSENKTLINEIAEKGGNTLAQVLTAWGLRGGYVLPKSSSNPKRIEINSPEIISD	286	
Db	225 --TDAPLKEPVIEAKNNVQPQGHVVIWHORGYVVLPKSVNPDRITNRKIFTLST	282	
Qy	287 ADPFAINAVAK 297		
Db	283 EDFPAINNISK 293		
RESULT 2	YDG7-SCHPO	STANDARD;	PRT; 321 AA.
AC	Q10494;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DE	Probable oxidoreductase C26F1.07 in chromosome I (EC 1.-.-.-).		
GN	SPAC26F1.07.		
OS	Schizosaccharomyces pombe (Fission yeast).		
OC	Eukaryota; Ascomycota; Schizosaccharomyces;		
OC	Schizosaccharomycetales; Schizosaccharomyctaceae;		
OX	NCBI_TAXID=4896;		
RN	SEQUENCE FROM N.A.		
RC	SEQUENCE=72;		
RX	MEDLINE:2184401; PubMed:118593360;		
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,		
RA	Scourgos J.J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,		
RA	Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,		
RA	Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,		
RA	Gentiles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,		
RA	KIANSKGEGGTGATIAVSNAITRGTSVPIKSVNQEPIKSNFK1YPLIKEDMDDEINSI-G 297		
Db	240 KIANSKGEGGTGATIAVSNAITRGTSVPIKSVNQEPIKSNFK1YPLIKEDMDDEINSI-G 297		
Qy	241 EIAE-KG-GNTLAAQVLYANGLRRGVLPKSSSNPKRIEINSPEIISDADFEAINAVAKG	298	
Db	298 IARRF 303		

Protein search results					
May 28, 2004, 13:59:56 ; Search time 1.8 Seconds (without alignments)					
940.156 Million cell updates/sis updates/sis					
MSNGKFTLNSGVKIEGVGF.....MKDTFFGYDVWVBBETAKNLUSA 325					
BLOSUM62					
Gapop 10.0 , Gapext 0.5					
Total number of hits satisfying chosen parameters: 141681					
Database : SwissProt_42 : *					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.					
SUMMARIES					
Result No.	Score	Query Match Length	DB ID	Description	
1	574.5	33.2	312	1 GCY YEAST	P14065 saccharomyces cerevisiae
2	568	32.8	321	1 YDG7_SCHPO	Q10494 schizosaccharomyces pombe
3	559.5	32.3	322	1 YDIX_SPOSA	P27800 sporobolus
4	555	32.1	312	1 YPRL1 YEAST	Q12458 saccharomyces cerevisiae
5	54.5	31.6	344	1 APAL_YEAST	P38115 saccharomyces cerevisiae
6	503	29.1	315	1 ALDI_MOUSE	P21300 mus musculus
7	501.5	29.0	310	1 S6PD_MALDO	P28475 malus domestica
8	499.5	28.9	315	1 ALDR_RAT	P07943 rattus norvegicus
9	489.5	28.3	284	1 P100_LBIMA	P22045 leishmania major
10	488.5	28.2	315	1 ALDR_PIG	P80276 sus scrofa
11	488.5	28.2	315	1 ALDR_RABIT	P15122 oryctolagus cuniculus
12	487.5	28.2	315	1 ALDR_MOUSE	P45376 mus musculus
13	487	28.1	316	1 AKBA_HUMAN	O60218 homo sapiens
14	486.5	28.1	315	1 ALDR_HUMAN	P15121 homo sapiens
15	486	28.1	318	1 XYL1_PACTA	P70694 pachysolen tunicum
16	486	28.1	323	1 DBB5_MOUSE	P45377 mus musculus
17	485	28.0	315	1 FLD2_MOUSE	Q9d211 mus musculus
18	480.5	27.8	301	1 AKEL1_MOUSE	P16116 bos taurus
19	478	27.6	315	1 ALDR_BOVIN	O9j167 mus musculus
20	478	27.6	324	1 AKAC1_MOUSE	P14550 homo sapiens
21	477	27.6	324	1 AKAL1_HUMAN	P02199 pseudomonas aeruginosa
22	475.5	27.5	323	1 PB2R_RABBIT	P80508 oryctolagus cuniculus
23	473	27.3	324	1 AKAL1_RAT	P51635 rattus norvegicus
24	467	27.0	324	1 AKAL1_DOG	P50578 sus scrofa
25	461	26.6	326	1 AKD1_HUMAN	P51857 homo sapiens
26	457	26.4	323	1 AKC2_HUMAN	P52897 bos taurus
27	456	26.3	295	1 MORA_SEPEU	P52895 pseudomonas fluorescens
28	454	26.2	323	1 AKC1_HUMAN	P04828 haldo-
29	450.5	26.0	323	1 EGFS_BOVIN	P05980 bos taurus
30	447.5	25.9	323	1 AKC3_HUMAN	P42330 haldo-
31	445.5	25.7	323	1 EGFR_BOVIN	P52897 bos taurus
32	445	25.6	323	1 EGFR_BOVIN	P52895 saccharomyces cerevisiae
33	443.5	25.6	323	1 DDX4_BOVIN	P52893 bos taurus

34	439	25.4	329	1	XYL1_KLULA	P49378	kluuyveromyc
35	438.5	25.3	323	1	AKC4_HUMAN	P17516	homo sapien
36	437	25.2	326	1	AKDI_RAT	P31210	rattus norv
37	435.5	25.2	323	1	CRO_EBANCA	P17564	rana catesbe
38	432.5	25.0	323	1	CRO_RANTE	P02532	rana tempor
39	432	25.0	320	1	ALDR_HORVU	P23901	hordium vul
40	430.5	24.9	323	1	PE2R_RAT	P51632	rattus norv
41	420.5	24.3	318	1	XYL1_EICST	P31867	pitchia stip
42	419.5	24.2	322	1	DIDH_RAT	P23457	rattus norv
43	414.5	23.9	321	1	DTDH_NUMCMJ	Q01213	microm cuced
44	414	23.9	321	1	DKGA_ECOLI	Q46857	escherichia
45	413	23.9	282	1	YJ66 YEAST	P47137	sacharamyc

ALIGNMENTS

searched: 141681 seqs, 52070155 residues
 total number of hits satisfying chosen parameters: 141681
 minimum DB seq length: 0
 maximum DB seq length: 200000000
 post-processing: Minimum Match 0%

卷之三

Pred. No. 15 The number of results pre-

SUMMARIES							Description
Result No.	Score	Query Match	Length	DB	ID	%	
1	574.5	33.2	312	1	GCY YEAST	1	P14665 sacchar
2	568	32.8	321	1	YDG7 SCHPO	1	Q1094 schitos
3	559.5	32.3	322	1	ALDX_SPOSA	1	P2700 sporobo
4	555	32.1	312	1	YBRI YEAST	1	Q1258 sacchar
5	547.5	31.6	344	1	ARAI YEAST	1	P14615 sacchar
6	529	29.1	315	1	ALDI_MOUSE	1	P21300 mus mus
7	501.5	29.0	310	1	SGD_MALDO	1	P2875 malus d
8	499	28.9	315	1	ALDR RAT	1	P07943 rattus
9	489.5	28.3	284	1	P100 LEIMA	1	P22454 leishma
10	488.5	28.2	315	1	ALDR_PIG	1	P8076 sus scir
11	488.5	28.2	315	1	ALDR_RABBIT	1	P12272 oryctol
12	487	28.2	315	1	ALDR_MOUSE	1	P45376 mus mus
13	487	28.1	316	1	AKBA_HUMAN	1	Q6018 homo sa
14	486.5	28.1	315	1	ALDR_HUMAN	1	P06126 homo sa
15	486	28.1	318	1	XXL1_FACTA	1	P78336 pachyso
16	486	28.1	323	1	DHB5_MOUSE	1	P70944 mus mus
17	485	28.0	315	1	ALD2_MOUSE	1	P71775 mus mus
18	480.5	27.8	201	1	AKE1_MOUSE	1	Q9dt1 mus mus
19	478	27.6	315	1	ALDR_BOVIN	1	P16116 bos taur
20	478	27.6	324	1	AKAI_MOUSE	1	Q9j16 homo mus
21	477	27.6	324	1	AKA1_HUMAN	1	P14550 homo sa
22	478	27.5	323	1	P22R_RABIT	1	P80505 oryctol
23	473	27.3	324	1	AKAI_RAT	1	P51315 rattus
24	467	27.0	324	1	AKA1_PIG	1	P05078 sus scir
25	461	26.6	326	1	AKD1_HUMAN	1	P51557 homo sa
26	457	26.4	323	1	AKC2_HUMAN	1	P52975 haldo
27	456	26.3	295	1	MORA_PSEPU	1	P02198 pseudom
28	454	26.2	323	1	AKC1_HUMAN	1	Q04428 haldo
29	450.5	26.0	323	1	PGFS_BOVIN	1	P05809 bos taur
30	447.5	25.9	323	1	AKC3_HUMAN	1	P42330 haldo
31	445.5	25.7	323	1	PGF2_BOVIN	1	P5297 bos taur
32	445	25.7	327	1	GRB3 YEAST	1	P52915 sacchar
33	443.5	25.7	323	1	DRY_BOVIN	1	P52926 bos taur

RESULT 1					
GCY - YEAST	STANDARD;	PRT;	312 AA.		
P14765;					
AC	01-JAN-1990	(Rel. 13, Created)			
DT	01-JAN-1990	(Rel. 13, Last sequence update)			
DT	28-FEB-2003	(Rel. 41; Last annotation update)			
DE	GCY protein (EC 1.1.1.-)				
GN	GCY or GCY OR YOR120W OR 031567 OR YOR329W.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomyces; Saccharomycetidae; Saccharomyces; Saccharomyces.				
OC					

RP SEQUENCE FROM N.A.
 RP STRAIN=D73-10B;
 RC MEDLINE=89005653; PubMed=2901985;
 RX
 RA Oeschner U., Magdolen V., Bandlow W.;
 RT "A nuclear yeast gene (Gcy) encodes a polypeptide with high homology
 RT to a vertebrate eye lens protein.";
 RL PNAS Lett. 238:123-128(1988).
 [12]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=S228C / FY1679;
 RX MEDLINE=97050020; PubMed=8904341;
 RA Wiemann S., Rechmann S., Benes V., Voss H., Schwager C., Vlcek C.,
 RA Stegemann J., Zimmermann J., Erflie H., Paces V., Ansorge W.;
 RT "Sequencing and analysis of 51 kb on the right arm of chromosome XV
 RT from *Saccharomyces cerevisiae* reveals 30 open reading frames.";
 RL yeast 12:281-288(1996).
 [3]
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9734368; PubMed=9200815;
 RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoro C.,
 RA Schwager C., Paces V., Sander C., Ansorge W.;
 RA "DNA sequencing and analysis of 130 kb from yeast chromosome XV.";
 RL yeast 13:655-672(1997).
 [4]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25657;
 RA Bandlow W.;
 RL Submitted (MAR-1986) to the EMBL/GenBank/DDBJ databases.
 CC -|- FUNCTION: Unknown function which seems to be not essential.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE FAMILY.
 CC -|- VERY SIMILAR TO YEAST YPR1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no
 CC modified and this statement is not removed. Usage by and for commer-
 CC cial entities requires a license agreement (See <http://www.isb-sib.ch/annex>
 CC or send an email to license@isb-sib.ch).
 CC
 EMBL: X13228: CAA31615:1; -

Db 3 TYOLNNNGVEIPVLFGGTPEKADGE-EAYRAYLEALKAGYRHDIAIYONEESVGAIKD 61
 Qy 66 FLKENPSVYKREDIFVCTKWMHLHRYEDVLNSIDSILKRLGIDYDMMPLFLWMPPIAEKNG 125
 Db 62 ---- SCVPREEMFVTIKLWNSSQTYEQTRQALESKTEKLGLDYLILHWP----- 109
 Qy 126 QGPBKIGPGKVKYLKDLTENPEP-----TWRAKEKIYEDRKARSIGVSNTW 172
 Db 110 -----NPXPLRENDAWTRNAEWTRAMEDLYQEGKRAIGVSNFL 149
 Qy 173 IADLEKMSKPAKMPHANQIEHPFLNEELVQCYFSKVNIMPVAYSPLGSQNQVPTGER 232
 Db 150 PHFLDALIETATIVPAYNQVRLLAPGVYDQDVAYCREGKILLEAKGPFQ-----GE- 201
 Qy 233 VSENKTLINEIAEKGGNTLAQVLIJAWGLRGTVVLPKSSNPKRIESNFK--SIELSDADFE 290
 Db 202 LFDSKQVQEIAAHNGKSVAQIALAWSLAEGLPLPKSVTTSRIQANLDCFGELSHEERE 261
 Qy 291 AINVA 296
 Db 262 TIKTIA 267

RESULT 15

C98038 conserved hypothetical protein spr1332 [Imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: 98038
 R:Hoskins, J.A.; Alborn, Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; Dehoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Leffkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J.:Bacteriol., 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young, Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: C98038
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-280 <KOR>
 A:Cross-references: GB:AE007317; PIDN:AAL00136.1; PID:915458977; GSPDB:GN00174
 C:Genetics:
 C:Gene: spr1332
 C:Supertamily: aldehyde reductase

Query Match 29.8%; Score 516; DB 2; Length 280;
 Best Local Similarity 35.9%; Pred. No. 7.8e-33;
 Matches 110; Conservative 57; Mismatches 83; Indels 56; Gaps 7;
 C:Accession: 98038

Db 6 TFTLSNGYKRPVGFITFASEGKGETYATTALKTGYRHDCAVYYLNGEVGEGIRD 65
 3 TYOLNNNGVEIPVLFGGTPEKADGE-EAYRAYLEALKAGYRHDIAIYONEESVGAIKD 61
 Qy 66 FLKENPSVYKREDIFVCTKWMHLHRYEDVLNSIDSILKRLGIDYDMMPLFLWMPPIAEKNG 125
 Db 62 ---- SCVPREEMFVTIKLWNSSQTYEQTRQALESKTEKLGLDYLILHWP----- 109
 Qy 126 QGPBKIGPGKVKYLKDLTENPEP-----TWRAKEKIYEDRKARSIGVSNTW 172
 Db 110 -----NPXPLRENDAWTRNAEWTRAMEDLYQEGKRAIGVSNFL 149
 Qy 173 IADLEKMSKPAKMPHANQIEHPFLNEELVQCYFSKVNIMPVAYSPLGSQNQVPTGER 232
 Db 150 PHFLDALIETATIVPAYNQVRLLAPGVYDQDVAYCREGKILLEAKGPFQ-----GE- 201
 Qy 233 VSENKTLINEIAEKGGNTLAQVLIJAWGLRGTVVLPKSSNPKRIESNFK--SIELSDADFE 290
 Db 202 LFDSKQVQEIAAHNGKSVAQIALAWSLAEGLPLPKSVTTSRIQANLDCFGELSHEERE 261
 Qy 291 AINVA 296
 Db 262 TIKTIA 267

ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.-P.X.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24487
A;Accession: T48188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <BEV>
A;Cross-references: EMBL:AL161946
A;Experimental source: cultivar Columbia; BAC clone F7A7
C;Genetics:
C;Map Position: 5
A;Introns: 48/3; 65/3; 90/2; 119/3; 188/3; 210/3; 291/3
A;Note: F7A7.190
C;Superfamily: aldehyde reductase

Query Match Score 521; DB 2; Length 320;
Best Local Similarity 35.6%; Pred. No. 3.8e-33;
Matches 117; Conservative 73; Mismatches 105; Indels 34; Gaps 10;

Qy 3 NGKTFILSNGYKIPGVGFCTFASEGSKGETYTAVTTALKTGYRLDCAWYVNLNEGEVGEG 62
Db 12 NMESFRILLSGKIPAVGLGTWRS-GSA-AHAWVTAIVECGYRHTDWEYGDREVGOG 69

RESULT 12

T09670 abscisic acid activated protein - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Accession: T09670
R;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
R;Deil, M.; Pavletova, S.; Oberschall, A.; Meszaros, T.; Misholczi, P.; Torkok, K.; Dudit submitted to the EMBL Data Library, April 1996
A;Description: Calmodulin inhibitors enhance the expression of abscisic acid-activated g
A;Reference number: Z16612
A;Accession: T09670
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-313 <DEA>
A;Cross-references: EMBL:X97606
A;Experimental source: cultivar Regen S; strain RA3
C;Superfamily: aldehyde reductase

Query Match Score 523.5; DB 2; Length 313;
Best Local Similarity 37.2%; Pred. No. 2.4e-33;
Matches 123; Conservative 61; Mismatches 106; Indels 41; Gaps 10;

Qy 5 KTFPLSNGKIPGVGFCTFASEGSKGETYTAVTTALKTGYRLDCAWYVNLNEGEVGEGIR 64
Db 6 KFFQLNTGAKIPPSVLTGQAE- -PGVVAKAVTAVOYGRHEDCAEAYKNQSEIGSALK 63

Qy 65 DFLQENPSTVSKREDIFVCTKWNHJRYEVDLWSTDDSLRKLDYVDMFVWPIAAEN 124
Db 64 K-LCEDGVVKREELWITSCLWCSDHHPEDVPKAQDTLNLDQDLYLHNPVSM-KR 121

Qy 125 GGQEPKIGPGKYYVILKDLTENPE----PTTRAMEKTYEDRKARSICSVNNTIADLEM 179
Db 122 GTGE----FGENLHDADISTWKLGAQHDSGKAKAIGNSNSTKQJDL 168

Qy 180 SKFAKUMPHANQIEIHPFLPNEELVQYCPFSKINMPVAYAPLGSONQVPTGEVSENKT 239
Db 169 LDVARPPATNQVEHPGQAKUAHFCSKGHLISGYPGLGSVLSKSD--ILKNEPVV 225

Qy 240 NEIAKGGNTLAQVLJANGRGGYVVLPKSSNPKRIESNFKSIELSAD----FEAI- 292
Db 226 KEIAERKLGKTPGQVALWQGAGSVLPKSTNEARIKQNDVYDWSPIDLFPKFSEIKQ 285

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <RUR>
A;Cross-references: GB:AB005672; PIDN:AAK75572.1; PMID:21357209; PMID:11463916
C;Genetics:
C;Map Position: 5
A;Introns: 48/3; 65/3; 90/2; 119/3; 188/3; 210/3; 291/3
A;Note: F7A7.190
C;Superfamily: aldehyde reductase

RESULT 13

T48188 aldoose reductase-like protein - Arabidopsis thaliana
N;Alternate names: protein F7A7.190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
C;Accession: T48188
R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De

Query Match Score 516; DB 2; Length 280;
Best Local Similarity 35.9%; Pred. No. 7.8e-33;
Matches 110; Conservative 57; Mismatches 83; Indels 56; Gaps 7;

Qy 6 TFTLSNGVKIPGVGFCTFASEGSKGETYTAVTTALKTGYRLDCAWYVNLNEGEVGIRD 65

aldo-keto reductase (EC 1.1.1.1) YPR1 - yeast (*Saccharomyces cerevisiae*)
 N;Alternate name: protein YDR368w
 C;Species: *Saccharomyces cerevisiae*
 C;Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
 C;Accession: S61163; S61563; JC447; PC4473
 R;Ding, H.
 Submitted to the EMBL Data Library, June 1995
 A;Description: The sequence of *S. cerevisiae* cosmid 9481.
 A;Reference number: S61159
 A;Accession: S61163
 A;Molecule type: DNA
 A;Cross-references: EMBL:U28373; PIDN:9849184; PIDN:AAB64804.1; PMID:849189; MIPS:YDR368w
 R;Miosga, T.; Juhnke, H.; Sterkelaar, C.; Zimmermann, F.K.
 submitted to the EMBL Data Library, July 1994
 A;Description: A new yeast gene with homology to the aldo-keto reductase protein family.
 A;Reference number: S61562
 A;Accession: S61563
 A;Molecule type: DNA
 A;Residues: 1-312 <DIN>
 A;Cross-references: EMBL:X00642; PIDN:9103496; PMID:9103498
 A;Experimental source: Strain M5
 R;Nakamura, K.; Kondo, S.; Kawai, Y.; Nakajima, N.; Ohno, A.
 Biosci. Biotechnol. Biochem. 61, 375-377, 1997
 A;Title: Amino acid sequence and characterization of aldo-keto reductase from bakers' yeast
 A;Reference number: JCS457; PMID:9722041; PMID:9059981
 A;Accession: JC5457
 A;Molecule type: Protein
 A;Residues: 2-36;45-50;68-96; 'X';1124-136;178-227;287-293;299-304 <NA2>
 A;Gene: SGD:YPR1
 A;Cross-references: MIPS:YDR368w; SGD:S0002776
 A;Map position: 4R
 C;Complex: monomer
 C;Function:
 A;Description: NADPH-dependent keto ester reductase; catalyzes the reduction of aldehydes to their corresponding alcohols
 C;Superfamily: aldehyde reductase
 C;Keywords: monomer; NADP; oxidoreductase
 P;2/312/Product: aldo-keto reductase #status experimental <MAT>

Query Match Score 555; DB 2; Length 312;
 Best Local Similarity 40.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 Db 7 NSSATLKLNGTASPIVLFQFTGTRSVNDNG-YHSVIAKAGRHRDAAVINEEVGR 64

Query 2 SNGKTFTLNGNGVKIPGVGKTFASEGSKGEVTAVTALKTGVRHLDCAWYLNCGEVGE 61
 Best Local Similarity 32.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 Db 62 GIRDFLKENPSVKEFIDFVCTWNHLHYREDLTLWSIDDSLRGLDYDMLVHMPVIA 121

Query 65 AIKD-----SGVPREEIFITKGTEB-DPBAALNSLRLGLDYDMLVHMPVPL 117
 Best Local Similarity 30.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 Db 122 EKNGQGPKEKGPKDGYKVILKDIIE-----NPEFTWAMEKTYEDRKARSTGVSNN 171

Query 118 KTDREV-----TDGVNLCPITLBDGTVDIDTKEWNFIKTVNQELRTGKTKAVGSNF 171
 Best Local Similarity 28.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 Db 172 TIADLEKM-SKEFAKMPHANOIEIHPFLPNEELVQYCFSKRNIMPWAVSPLGSONQVPTT 229

Query 172 SINNTEKLNGTASPIVLFQFTGTRSVNDNG-YHSVIAKAGRHRDAAVINEEVGR 64
 Best Local Similarity 28.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 Db 172 SNNNTEKLNGTASPIVLFQFTGTRSVNDNG-YHSVIAKAGRHRDAAVINEEVGR 64

Query 230 GERYSENKTLINEAEGGNTLAQVLIANGLRRGGYVVLPKSNPKNRISNPKSIELSDADF 289
 Best Local Similarity 28.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;
 Db 231 KE-----QAIIDMAKKHGVEPAQIISMSIQRGYYVLAKSNPNERTVSNFKTFPLPDDF 285

Query 290 EAIVAVAKGRHF-RFVNMK 307
 Best Local Similarity 28.1%; Pred. No. 8.1e-36; Gaps 9;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 9;

Db 286 KTISNLSKYHGTXRVVDMK 304
 Best Local Similarity 28.1%; Pred. No. 8.1e-36; Gaps 9;

RESULT 5
 S61163 probable aldehyde reductase (EC 1.1.1.1) - *Synechocystis* sp. (strain PCC 6803)
 C;Species: *Synechocystis* sp.
 A;Variety: PCC 6803
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C;Accession: S76143
 R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, B.; Miyajima, N.; O., K.; Okamura, S.; Shimojo, S.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
 A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
 A;Reference number: S76143
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Cross-references: EMBL:D90914; GB:AB001339; PIDN:BA018402.1; PMID:8905231
 A;Cross-references: EMBL:D90914; GB:AB001339; PIDN:BA018402.1; PMID:97061201; PMID:8905231
 A;Accession: S76143
 A;Start codon: GTG
 C;Superfamily: aldehyde reductase
 C;Keywords: oxidoreductase

Query Match Score 553; DB 1; Length 327;
 Best Local Similarity 38.1%; Pred. No. 1.2e-35;
 Matches 117; Conservative 69; Mismatches 105; Indels 16; Gaps 6;

Query 5 KTFITLSNGVKIPGVGKTFASEGSKGEVTAVTALKTGVRHLDCAWYLNCGEVGE 64
 Db 11 KYFPLNSGEQIPALGLGTMKS--SPQVYQGAVEQALDGYRHDCAAYTGNEAEIGTLA 68

Query 65 DFLKENPSVKEFIDFVCTWNHLHYREDLTLWSIDDSLRGLGDYVDFLWHPVIAEKN 124
 Db 69 NAFTG-VVKREELWTSKLWSNAHHPDAVLPALEKTQDGLDYLILHKPVVI-- 124

Query 125 QGQEPKIG-PDGKVTLKCOLLTENBPPTKMEKTYEDRKARSIGCVSNNTIADLBKMSFA 183
 Db 125 ---QDDVGFPESGDQLLPTPASLEGTVNLERAVDVGICHHGSNSNLSKLEMULSMA 181

Query 184 KVMPHANQQTBIHPFLPNEELVQYCFSKNIMPWAVSPLGSONQVPTTGERSEENKL--- 239
 Db 182 RIPPAVNVQELHPYHQQSDFLTAFNSQNLTAVSPLGSDGR-PAAFQQAEPKLITDPV 240

Query 240 -NEIAEKGNTLQLAQVLIANGLRRGGYVVLPKSSNPKBISNEFSKIELSDADFEAINAVAKG 298
 Db 241 INGIAEQQGSAAQVLLAVAIQRTVTPKSVNPERLBNLRASTITDSENKAIALD 300

Query Match Score 305; DB 2; Length 305;
 Best Local Similarity 30.1%; Pred. No. 8.1e-36; Gaps 0;
 Matches 128; Conservative 60; Mismatches 97; Indels 34; Gaps 0;
 Db 301 RHYRVS 307

RESULT 6
 T39168 probable oxido-reductase [imported] - fission yeast (*Schizosaccharomyces pombe*)
 C;Species: Schizosaccharomyces pombe
 C;Accession: T50378; T3169
 R;Connor, R.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, February 1998
 A;Reference number: Z21832
 A;Accession: T50378
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-325 <CO2>
 A;Cross-references: EMBL:AL021815; PIDN:CAA16997.1; GSPDB:GN00067; SPDB:SPBC8B4.04
 A;Experimental source: strain 972h-; cosmid C8B4
 C;Genetics:
 A;Map position: 2
 C;Superfamily: aldehyde reductase

A; Map position: 15R	Qy 241 EIAE-KG-GNTLAQVLIJAWGLRGGYVLPKSSNPKRIESNPKSIELSDADFEAINAVARG 298
C; Keyword: aldehyde reductase	Db 240 KIAKSSEGTGATGATTAVSWATRGSVIPSVNTERIKSNFKYPLTKDMDENSI--G 297
C; Keyword: oxidoreductase	
Query Match 33.2%; Score 574.5; DB 1; Length 312;	
Best Local Similarity 40.8%; Pred. No. 2.4e-37;	Qy 299 RHFRF 303
Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;	Db 298 IRARF 302
Qy 5 KTFTUSNGKRPQGCFGTIPASEGSKGETTAYTALKTGYRHLDCAWYLYNEGEVGGR 64	RESULT 3
Db 10 KILSINTGAQIPQICLGTQSX- ENDAYKAVALTALKTGYRHLDTAQNEDQVGAIK 67	S77113
Qy 65 DFLKENPSVSKREDIFVCTKWNHHLHYEDVLSIDSRLKGLDYDVMFLYHMPIAEKN 124	aldehyde reductase (NADPH) (EC 1.1.1.-) - fungus (Sporidiobolus salmonicolor)
Db 68 D----SGVPREEFVFTKLWCTGSHPEV-- ADQSLRKLGKDYDVLMLHMPARL-- 117	N; Alternative names: NADPH-dependent aldehyde reductase
Qy 125 GOGERKIGPDKKYTLKDLTENPE-----PTWRAKEKIYEDRKARSTGV 168	C; Species: Sporidiobolus salmonicolor, Sporobolomyces salmonicolor
Db 118 -----DPAYLKNEEDLISVTPVKDGRAVDITNNWKTBLMQBLPKTRKTRAVGV 168	C; Date: 29-Jan-1998 #sequence revision 06-Feb-1998 #text_change 11-Jun-1999
Qy 169 SNWTIADLIERM-SKFKAKMMPHANOIETHPFLNEELVQYCFSKINPVAYSPGSQNC 226	C; Accession: S77113; S23331
Db 169 SNFSNLNLKDLASOGNKTPAANQVEIPLLPODELINFCCKSGKIVVEAYSPLGS --- 224	R; Kit: Kita, K.; Matsuzaki, K.; Hashimoto, T.; Yanase, H.; Kato, N.; Chung, M.C.; Kataoka, M.; Miyoshi, T.; Shimizu, S.; Yamada, H.
Qy 227 PTGEBRVSENKNTLBAIAEGGNTIAQVLIAWGLERGGYVLPKSSNPKRIESNPKSIELSD 286	A; Cross-references: EMBL:U26463; NID:91142697; PID:AAB17362.1; PID:91142698
Db 225 -TDAPLXKPVILEAKONVQEPHVVWSHVQGYVWLPKSYVNPDRTKTRKIFTLST 282	A; Experimental source: strain AKU 4429
Qy 287 ADFEAINAVAK 297	R; Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
Db 283 EDFEAINNISK 293	D; Submitted to the EMBL Data Library, April 1996
Qy 298 T38413	A; Reference number: Z21792
Db 299 A; Status: preliminary; translated from GB/EMBL/DDBJ	A; Accession: T38413
Qy 300 A; Molecule type: DNA	A; Cross-references: EMBL:Z73100; PIDN:CAA97364.1; GSFD:GN00066; SPDB:SPAC26F1.07
Db 301 C; Genetics:	A; Experimental source: strain 972h-; cosmid c26f1
Qy 302 A; Gene: SPDB:SPAC26F1_07	A; Map position: 1
Db 303 C; Map position: 1	C; Map position: 1
Qy 304 T38413	C; Position: aldehyde reductase
Db 305 T38413	Query Match 32.8%; Score 568; DB 2; Length 321;
Qy 306 Best Local Similarity 43.0%; Pred. No. 8.2e-37;	Best Local Similarity 39.3%; Pred. No. 3.8e-36;
Db 307 Matches 131; Conservative 59; Mismatches 89; Indels 26; Gaps 10;	Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 308 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 309 Db 310	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 311 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 312 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 313 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 314 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 315 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 316 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 317 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 318 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 319 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 320 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 321 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 322 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 323 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 324 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 325 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 326 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 327 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 328 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 329 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 330 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 331 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 332 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 333 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 334 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 335 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 336 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 337 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 338 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 339 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 340 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 341 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 342 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 343 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 344 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 345 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 346 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 347 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 348 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 349 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 350 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 351 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 352 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 353 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 354 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 355 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 356 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 357 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 358 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 359 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 360 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 361 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 362 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 363 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 364 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 365 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 366 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 367 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 368 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 369 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 370 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 371 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 372 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 373 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 374 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 375 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 376 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 377 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 378 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 379 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 380 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 381 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 382 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 383 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 384 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 385 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 386 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 387 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 388 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 389 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 390 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 391 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 392 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 393 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 394 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 395 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 396 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 397 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 398 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 399 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 400 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 401 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 402 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 403 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 404 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 405 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 406 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 407 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 408 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 409 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 410 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 411 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 412 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 413 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 414 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 415 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 416 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 417 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 418 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 419 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 420 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 421 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 422 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 423 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 424 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 425 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 426 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 427 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 428 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 429 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 430 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 431 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 432 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 433 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 434 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 435 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 436 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 437 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 438 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 439 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 440 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 441 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 442 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 443 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 444 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 445 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 446 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 447 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 448 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 449 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 450 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 451 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 452 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 453 T38413	Query Match 32.3%; Score 559.5; DB 2; Length 323;
Db 454 T38413	Best Local Similarity 39.3%; Pred. No. 3.8e-36; Matches 127; Conservative 60; Mismatches 119; Indels

XX WO200210425-A2.

XX PPN

XX PD 07-FEB-2002.

XX PPF 02-AUG-2001; 2001WO-GB003485.

XX PPR 02-AUG-2000; 2000US-00630983.

XX (BIOPOL-) BIOPOLÒ SCARL.

XX (WHALE/) WHALLEY K.

XX PPI Porro D., Sauer M;

XX DR WPI: 2002-217125/27.

XX N-PSSDB; ABK10132.

XX PPT Generating ascorbic acid or its salt, involves culturing yeast capable of converting ascorbic acid precursor into ascorbic acid in medium comprising ascorbic acid precursor, and isolating ascorbic acid.

XX PPS Claim 12; Page 90-91; 95pp; English.

XX The invention relates to generating ascorbic acid or its salt, involves obtaining a recombinant yeast capable of converting an ascorbic acid precursor into ascorbic acid, culturing the recombinant yeast in a medium comprising an ascorbic acid precursor, thus forming ascorbic acid, and isolating the ascorbic acid. Also include are stabilising ascorbic acid or its salt in a medium, by culturing a yeast in a medium comprising ascorbic acid or its salt and a recombinant yeast functioning as a transformed with a coding region encoding a protein having an enzyme activity selected from L-galactose dehydrogenase (LGDH), L-galactono-1, 4-lactone dehydrogenase (LDH), D-arabinose dehydrogenase (ARA), D-arabinono-1,4-lactone oxidase (ALO), L-gulono-1,4-lactone oxidase (GLO) and aldonolactonase (AL) activity, where the recombinant yeast is capable of converting to ascorbic acid at least about 25% ascorbic acid precursor or is capable of producing at least 20 mg ascorbic acid/L medium, when

RESULT
 ABP53552 ID ABP53552 standard; protein; 328 AA.
 XX AC ABP53552;
 XX AC ABP53552;
 DT 16-DEC-2002 (first entry)
 XX DE T. megachiliensis erythrose reductase type I SEQ ID NO:3.
 XX DE T. megachiliensis erythrose reductase type I SEQ ID NO:3.
 XX KW Erythrose reductase; enzyme; erythritol; D-erythrose.
 XX OS Trichosporonoides megachiliensis.
 XX PN EP1221478-A2.
 XX PD 10-JUL-2002.
 XX PF 04-JAN-2002; 2002EP-000000321.
 XX PR 09-JAN-2001; 2001JP-000001294.
 XX PR 18-DEC-2001; 2001JP-00384357.
 XX PA (NORO) NAT FOOD RES INST MIN AGRIC.
 XX PA (NIKIM) NIKEN CHEM CO LTD.
 XX PI Ookura T, Kasumi T, Asaba E;
 XX DR WPI: 2002-659435/71.
 XX DR N-PSDB; AB082227.
 XX PT Novel protein having erythrose reductase activity such as erythrose
 XX PT reductase type I, II or III, useful for production of erythritol on an
 XX PT industrial scale.
 XX PS Claim 17; Page 24-26; 39pp; English.

RESULT 13

XX	WO200264766-A2.		
PN	ID ABG93050 standard; protein; 344 AA.		
XX	ID ABG93050;		
PD	XX		
XX	AC ABG93050;		
PF	XX		
21-DEC-2001; 2001WO-EP015398.	XX		
XX	DT 21-NOV-2002 (first entry)		
PR	XX		
22-DEC-2000; 2000EP-00870318.	XX		
PR	DE S. cerevisiae BAX-associated protein fragment SEQ ID 58.		
04-JAN-2001; 2001EP-00870002.	XX		
PR	XX		
09-JAN-2001; 2001EP-00870003.	KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;		
XX	KW vasotropin; vaccine; gene therapy; proliferative disorder; cancer;		
PA (JANS) JANSSEN PHARM NV.	KW apoptosis; fungal; yeast infection; autoimmune disease; ischaemia;		
XX	KW neurodegeneration; cell death.		
XX	XX		
Contreras RH, Eberhardt I, Luyten WHM, Reekmans RJ;	Saccharomyces cerevisiae.		
XX	OS XX		
DR N-PSDB; ABQ76369.	OS XX		
XX	PN WO200264766-A2.		
PT New isolated nucleic acid representing a synthetic BAX-gene, useful as	PN XX		
PT medicament for treating, preventing and/or alleviating yeast or fungal	PN XX		
PT infections or proliferative disorders, or for preventing apoptosis in	PN XX		
PT certain diseases.	PN XX		
PS Claim 36; Fig 2; 344pp; English.	PP 21-DEC-2001; 2001WO-EP015398.		
XX	XX		
This invention describes a novel nucleic acid representing a synthetic	XX	22-DEC-2000; 2000EP-00870318.	
Bax Gene. The Bax gene of the invention is useful for identifying Bax-	XX	PR 04-JAN-001; 2001EP-00870002.	
resistant yeast or fungi; identifying, or obtaining and identifying	XX	PR 09-JAN-2001; 2001EP-00870003.	
Candida spp. sequences that are differentially expressed in a pathway	XX	XX	
CC eventually leading to programmed cell death or identifying inhibitors or	PA (JANSSEN PHARM NV		
CC inhibitor sequences of Bax-induced cell death. The products of the	PA		
CC invention have cytostatic, fungicide, immunosuppressive and	PI Contreras RH, Eberhardt I, Luyten WHM, Reekmans RJ;		
CC vasotropin activity and can be used in vaccines or for gene therapy. The	XX		
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,	XX		
CC antisense molecules and antibodies are useful as medicaments or in	XX		
CC preparing a medicament for treating, preventing and/or alleviating such	XX		
CC diseases associated with yeast or fungi or proliferative disorders, such	XX		
CC as cancer, or for preventing apoptosis in certain diseases. The compounds	XX		
CC or polypeptides, or the genetically modified organism are useful for	XX		
CC preparing a medicament for modifying the endogenous flora of humans and	XX		
CC other mammals. The vaccine is useful for immunising against yeast or	XX		
CC fungal infections. Apoptosis-related diseases include autoimmune disease,	XX		
CC ischaemia, diseases related with viral infections or neurodegenerations.	XX		
This sequence represents a polypeptide associated with the Bax gene	XX		
described in the disclosure of the invention	XX		
Sequence 295 AA:	XX		
Query Match 31.8%; Score 550; DB 5; Length 295;	XX		
Best Local Similarity 41.9%; Pred. No. 4.4e-45;	XX		
Matches 125; Conservative 56; Mismatches 87; Indels 30; Gaps 8;	XX		
Qy 3 NGKTTPLNSGVKIPGVGFITPASEBSKGEGYATTALTYRHLDCAWYYLNEGEVCG 62	CC		
Db 8 NTDFELPNNGNKTPAVGLGTWQATNEDDRAVLAALKNGYKHIDTAIYGNEQVGKA 66	CC		
Qy 63 IRDFLKENPSKREDIFVCTKWNHRYEDVMSDIDSLKRIGLDYDMFLYHWPIAA 122	CC		
Db 67 IKD---SGVPREELFVTKLNADHK--NIEEALETSLLKGJLNVYLHWPASID 119	CC		
Qy 123 KNGQGPKIGPDGKVYLKDLTENBPPTWRAKENYED-RKARSIGVSNWTIADLERM- 179	CC		
Db 120 RSTN---KPTDFDVT-----DTVRGLQKTVNFKRKLERLIS 166	CC		
Qy 180 SKFAKOMPHANOIEITHPPLVNEELVOYCFSKNMPVAYSPLGSONQVPTGERVSENKL 239	CC		
Db 167 SEGVDPVPAVNOIEAHPPLTQPELYDYLKEKGIVLEAVSPLGS ----TNSPLFKNETI 220	CC		
Qy 240 NEIAEKGGNTIAQVLIAGWGRGYVVLPGSNPRIESNFKSIELSDADEALNAVAK 297	CC		
Db 221 VKIABRNQVGPQAVLVSAQRKTVLPLSVTESRVNLSKTFPLPSEEFETLNKLSE 278	CC		

Query Match 31.6%; Score 547.5; DB 5; Length 344;

Best Local Similarity 37.1%; Pred. No. 9.7e-45;	XX	
Matches 58; Mismatches 51; Indels 55; Gaps 8;	XX	
Qy 7 FTLSNGVKIPGVGFITPASEBSKGEGYATTALTYRHLDCAWYYLNEGEVCGIRDF 66	CC	
Db 25 FSLNNNGRPVPAQLVWSAQRTKTVLPLSVTESRVNLSKTFPLPSEEFETLNKLSE 84	CC	
Qy 67 LRNPNSPVKREDIFVCTKWNHRYEDVNLW----SIDDSSLKRIGLDYDMFLVWPIAA 121	CC	

PT Carbonyl reductase gene and protein - used in the production of R-gamma-
 PT substituted-beta-hydroxybutyrate, for use in drugs and agrochemicals.
 XX
 PS Claim 1; Page 9-10; 14pp; Japanese.

XX The present sequence is that of a carbonyl reductase (ALD) isolated from
 CC Sporobolomyces salmonicolor IF01038. The enzyme is used in the synthesis
 CC of an (R)-gamma-substituted-beta-hydroxybutyrate (GSBH) from a gamma-
 CC substituted acetooctate. The GSBH is useful as a raw material for the
 CC synthesis of drugs and agricultural chemicals. (Updated on:16-OCT-2003 to
 CC standardise OS field)

SQ Sequence 323 AA;

```

Query Match Score 559.5; DB 2; Length 323;
Best Local Similarity 39.3%; Pred. No. 5.8e-46;
Matches 127; Conservative 60; Mismatches 119; Indels 17; Gaps 6;
Qy 6 TFTLSNGVKIPGVGFGTASBGSKGETVTAVTALKTYRHLDCAYVNLNGEVGSGIRD 65
Db 4 MTILNTGASLELVYGTW-QAAGPEVGQKVAIEGYRHDLAKVYSNQEVGAAKE 61
Qy 66 FLKNEPNPSVERDFVCTKWNHLDHRYEDLWSSIDSURKLGLDYDMFLVWPIAEKG 125
Db 62 ---AGVKEREDLPIITSKLWNNSHRPEQVEPALDTLKEGLYEVDYLJLTPVAFPPEG 116
Qy 126 QGEPKIGPDGKXYILK-DLTENDEPTWRAAMEKLYEDRKARSGVSNTIALEKNSKFAK 184
Db 117 DITONLFKPANDGEVKLDLIEVSLVDTWKAMVLLDITGKVKAIVSNPDAKNDAI-EATG 176
Qy 185 VMPHANOETIHPPLPNEELVQYCFSKNMIVPQPLGSQNSQNPPTG-ERVSSENKTLINEIA 243
Db 177 VTPSVNQITERPHLQLQPELIAHHAKANTHTIATSPGLNN---TVGPPLLQHPEKTRIA 232
Qy 244 EKGENTLAQVLIANGLRLGGYVVLPKSSNPKRIESNFKSIELSDADFAINAVAKGHERF 303
Db 233 EKNCTPAPQVLIAVVGHSVLPKSTPSRICENFQVSQDVASKGEGSRRR 292
Qy 304 VNMKDFTFG---YDVWPEETARN 322
Db 293 YNIFCTYSFKWDINVFGEEDEKS 315

```

RESULT 11

AGG3564 standard; protein; 313 AA.
 ID AGG3564;

AC AC63564;

DT 15-OCT-2001 (first entry)

XX Amino acid sequence of a ketone reductase.

XX Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
 XX directed evolution; molecular breeding; ketone reductase;
 XX Saccharomyces cerevisiae.

OS WO2001155442-A2.

PN PN

XX

PD PD

XX

XX

XX

XX

XX

XX

XX

XX

DR DR

PT Designing synthetic nucleic acid sequences for improved amplification,
 PT expression in host cell, by comparing free energy of folding of a
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement.

XX PS Claim 6; Page 69-70; 117pp; English.

XX The present sequence represents a ketone reductase. The polynucleotide
 CC sequence was modified using using the method of the invention. The
 CC specific sequence describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with a
 CC different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotides, and comparing
 CC this with that of the original polynucleotide. The method is useful for
 CC developing nucleic acid sequences that enhance expression of the encoded
 CC protein in heterologous host. The design and preparation of the
 CC synthetic genes are used in application of gene shuffling, directed
 CC evolution and molecular breeding methods. The method allows expression of
 CC genes from various organisms such as mammals, plants, yeast, fungi and
 CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
 CC hosts at commercially viable levels, in particular proteins with low
 CC yield such as methionine gamma-lyase from P. putida.

XX SQ Sequence 313 AA;

Query Match	Score 555;	DB 4;	Length 313;
Best Local Similarity	40.1%;	Pred. No. 1.5e-45;	
Matches 128;	Conservative 60;	Mismatches 97;	Indels 34; Gaps 9;

Qy	2 SNGKFTPLSNGVKIPGVGFGTASBGSKGETVTAVTALKTYRHLDCAYVNLNGEVGE 61
Db	8 NSSSATKLNTGASIPVLGSTWRSVNDN--YHSVIALALKAGYRHLDAATYLNEEGCR 65
Qy	62 GIRDFLKENPSVKREDIFVCTKWNHLDHRYEDLWSSIDSURKLGLDYDMFLHWPJAA 121
Db	66 AIKD---SGVPREBIFITTKLNGTER--DPEAALIKSLRGLDYLMLHWPPL 118
Qy	122 EKNGQGEPKIGPDGKXYVILKDLTE-----NPEPTWRAAMEKLYEDRKARSIGVSNW 171
Db	119 KTDRV---TDGRNVLCIPLTEDGTVIDTDTKENNFIKTWLMQELPTKGTKXAVGVSNF 172

RESULT 12

ABG93303
 ID ABG93303 standard; protein; 295 AA.

XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX

XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX

Candida albicans.

OS

5 KTFTLNSGYKIPGVGFPTASEGSKGETTYAVTTALKGYRHLDCAWYLNNEGEVGEGIR 64
 Qy 227 PITGERVSENKTLINEARKGGNTLAQVLIANGIARRGVVILPKSSNPKRTESENKFELSD 286
 Db 225 --TDAPLKEPVLIEARKNNVYQPGHVVISWHQRGTWVLPKSVNPDRIKNRKTFLST 282
 Qy 287 ADPEAINAVAK 297
 Db 283 EDPEAINNSK 293

RESULT 7
 AAG63566 standard; protein; 313 AA.
 XX AAG63566;
 AC
 XX DT-10CT-2001 (first entry)
 DE Synthetic amino acid sequence of a yeast protein.
 XX Methionine gamma-lyase; mda gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO200155342-A2.
 XX 02-AUG-2001.
 PD XX
 PF 31-JAN-2001; 2001WO-US003166.
 PR XX
 PR 31-JAN-2000; 2000US-0049421.
 PR 08-DEC-2000; 2000US-00734237.
 PA (BIOC-) BIOCATALYTICS INC.
 PA
 PI Rozzell DJ, Bui P, Hua L;
 XX
 DR; 2001-483235/52.
 DR N-PSDB; AAHT4584.

Claim 6; Page 73-74; 117pp; English.

XX Designing synthetic nucleic acid sequences for improved amplification.
 PT expressing in host cell, by comparing free energy of folding of a
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement.
 XX
 PS The present sequence is a modified yeast protein. The polynucleotide
 CC sequence was modified using the method of the invention. The
 CC specification describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with a
 CC different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotides, and comparing
 CC this with that of the original polynucleotide. The method is useful for
 CC developing nucleic acid sequences that enhance expression of the encoded
 CC protein in a heterologous host. The design and preparation of the
 CC synthetic genes are used in application of gene shuffling, directed
 CC evolution and molecular breeding methods. The method allows expression of
 CC genes from various organisms such as mammals, plants, yeast, fungi and
 CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
 CC hosts at commercially viable levels, in particular proteins with low
 CC yield such as methionine gamma-lyase from *P. putida*

Sequence 313 AA;

Score 574.5; DB 4; Length 313;
 Best Local Similarity 40.8%; Pred. No. 1.9e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

Qy 11 KILSINTAQIPIOGLGTMOSK- -ENDAYKAVALTALKGYRHLDCAWYLNNEGEVGEGIR 68
 Db 65 DFLKENPSYKRREDIFVCFVKWNHLHRYEDVLWSIDDSKLRLGDIYDVMFLVIEWPIAAEKN 124
 Qy 69 D----SGVPREEFVTKLWCTQHHEPEV--ALDSQSLKLRLGDIYDVLKLYNHPARL-- 118
 Db 125 QGEGFKIGDGKTYILKDLTENP-----PTTRAMENKYEDEKARSIGV 168
 Qy 119 -----DPAYTKNDLSPVTPKRDGSRAVDITNNFIKWLMOQLPKTGKTRAVGV 169

Db 169 SNWTIADLEKM--SKPAKUMPHANQIETHPLPNEELQYCFSKNIMPVASPLGSQNQV 226
 Qy 170 SNSFSINNLXDLIASQGNKLTPAANQVEHTPLPQDELINFCKSKGIVVEAYSPLGS--- 225

Db 227 PTTGGERVSENKTUNEIAEKGGNTLAQVLTAWGLERGTIVLPLKSSNPKRIESNFKSIBLSD 286
 Qy 226 -TDAPLXEPVILEIAKNNVQGPQHVVTSWVQRGYVULPKSSVNPDIKTNRKIFTLST 283

Db 287 ADFEAINAVAK 297
 Qy 284 EDFEAINNSK 294

RESULT 8
 AAW2218 ID AAW29218 standard; protein; 312 AA.
 XX AAW29218;
 AC AAW29218;
 XX DT 02-MAR-1998 (first entry)
 XX DE S. cerevisiae uronate dehydrogenase variant R57E.
 XX KW uronate reductase; uronate dehydrogenase; *Saccharomyces cerevisiae*;
 KW synthesis; L-ascorbic acid; variant.
 XX Saccharomyces cerevisiae.
 XX Location/Qualifiers
 Key-Misc-difference 57
 FT /label= R57E
 FT /note= "wild type arginine replaced by glutamine"
 XX DE19604798-A1.
 XX 14-AUG-1997.
 XX 09-FEB-1996; 96DE-01004798.
 PA (HERB-) HERBSTREITH & FOX PEKTIN-FAB NEUENBUERG.
 XX PI Mattes R, Kulbe K;
 XX DR; 1997-403701/38.
 XX Recombinant uronate reductase and L-gulono-gamma-lactone oxidase - used
 PT in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 XX
 PS Claim 31; Page; 26pp; German.
 XX The present sequence represents a uronate dehydrogenase (reductase) R57E
 CC variant. The variant can be used for synthesis of L-ascorbic acid. DNA
 CC encoding L-gulono-gamma-lactone oxidase (GuOx) is also used in the
 CC production of L-ascorbic acid. The uronate dehydrogenase has higher
 CC affinity for substrate (D-glucuronic or galacturonic acids) and coenzymes
 CC than the similar enzyme from *Lipomyces starkeyi*, also better stability
 CC and greater yield when produced recombinantly. NB. This sequence was
 CC created using the wild type uronate dehydrogenase sequence given in
 CC Figure 3 of the specification

PA (BIOC-) BIOCATALYTICS INC.
 XX PI Rozzell DJ, Bui P, Hua L;
 DR WPI: 2001-483235/52.
 DR N-PSDB; AAH74583.
 XX PT Designing synthetic nucleic acid sequences for improved amplification,
 PT expression in host cell, by comparing free energy of folding of a codon
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement.
 XX PS Claim 6; Page 72-73; 117pp; English.
 XX CC The present sequence represents a yeast protein. The polynucleotide
 CC sequence was modified using using the method of the invention. The
 CC specification describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with a
 CC different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotide, and comparing
 CC this with that of the original polynucleotide. The method is useful for
 CC developing nucleic acid sequences that enhance expression of the encoded
 CC protein in a heterologous host. The design and preparation of the
 CC synthetic genes are used in application of gene shuffling, directed
 CC evolution and molecular breeding methods. The method allows expression of
 CC genes from various organisms such as mammals, plants, yeast, fungi and
 CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
 CC hosts at commercially viable levels, in particular proteins with low
 CC yield such as methionine gamma-lyase from *P. putida*
 XX SQ Sequence 312 AA;
 Query Match 33.2%; Score 574.5; DB 4; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.9e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;
 Qy 5 KTFPLSNGKIPGVGFGTASEGSKGETTAAVTALKTGYRHLDCAWXYLNEGEVGGR 64
 Db 10 KILSLNTGAQIPOIGITWOSK- ENDAYKAVLTALKDGYRHDTRAYRNNEQVQGAIK 67
 Qy 65 DFLKENPSVKREDIFVCTYKWNHLHRYEDYLWSDDSLKRGLDGYDVMNLVMPMIAEKN 124
 Db 68 D-----SGVREEEFVTTKLWCTQHPEPV--ALDOSLKRGLDGYDVLMLHMPARL-- 117
 Qy 125 GQGEPKIGPKGYVLLKDLTENP-----PTVRAAMEKTYEDERKARSIGV 168
 Db 118 -----DPAYTKNEDILSVPTXKDGSRAVDITNNNPKTTELMOELPKTGKTKAVG 168
 Qy 169 SNWTLADLERK--SKPAKYMPHANOIEIHPFLPNBELYCYCFSKRNIMPVAYSPLGSQNQV 226
 Db 169 SNSFSINLKLASSONKLPAANQVEIPLPDELLNFCKRSIGIVWEAYSPLGS--- 224
 Qy 227 PTTGGERVSЕНКЛНДИАЕКГНТЛАQVЛГЛРГҮҮЛВКПКССНПКРЕСНПКСИЕЛД 286
 Db 225 --TDAPLIKKEPVILAKNNVQPGHVVISWHORGYVYLPSKSVNPDRKTNKIFTLT 282
 Qy 287 ADFEAINAVAK 297
 Db 283 EDFEAINTNSIK 293
 RESULT 6
 ABG93198 standard; protein; 312 AA.
 XX ABG93198;
 XX ABG93198;
 DT 21-NOV-2002 (first entry)
 XX DE *S. cerevisiae* BAX-associated protein fragment SEQ ID 354.
 XX DE S. cerevisiae BAX-associated protein fragment SEQ ID 354.
 XX OS *Saccharomyces cerevisiae*.
 XX PN WO200264766-A2.
 XX PD 22-AUG-2002.
 XX PF 21-DEC-2001; 2001WO-EP015398.
 XX PR 22-DEC-2000; 2000EP-00870318.
 PR 04-JAN-2001; 2001EP-00870002.
 PR 09-JAN-2001; 2001EP-00870003.
 XX PA (JANC) JANSSEN PHARM NV.
 PI Contreras RH, Eberhardt I, Luyten WHL, Reekmans RJ;
 XX DR WPI; 2002-667002/71.
 XX DR N-PSDB; ABQ76464.
 XX PT New isolated nucleic acid representing a synthetic BAX gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.
 XX PS Claim 36; Fig 1; 344pp; English.
 XX SQ This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying Bax-
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cytostatic, fungicide, immunosuppressive, virucide and
 CC vasotropoc activity and can be used in vaccines or for gene therapy. The
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
 CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenous flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polypeptide associated with the Bax gene
 XX SQ Sequence 312 AA;
 Query Match 33.2%; Score 574.5; DB 5; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.9e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;
 Qy 5 KTFPLSNGKIPGVGFGTASEGSKGETTAAVTALKTGYRHLDCAWXYLNEGEVGGR 64
 Db 10 KILSLNTGAQIPOIGITWOSK- ENDAYKAVLTALKDGYRHDTRAYRNNEQVQGAIK 67
 Qy 65 KTFPLSNGKIPGVGFGTASEGSKGETTAAVTALKTGYRHLDCAWXYLNEGEVGGR 64
 Db 10 KILSLNTGAQIPOIGITWOSK- ENDAYKAVLTALKDGYRHDTRAYRNNEQVQGAIK 67
 Qy 65 DFLKENPSVKREDIFVCTYKWNHLHRYEDYLWSDDSLKRGLDGYDVMNLVMPMIAEKN 124
 Db 68 D-----SGVREEEFVTTKLWCTQHPEPV--ALDOSLKRGLDGYDVLMLHMPARL-- 117
 Qy 125 GQGEPKIGPKGYVLLKDLTENP-----PTVRAAMEKTYEDERKARSIGV 168
 Db 118 -----DPAYTKNEDILSVPTXKDGSRAVDITNNNPKTTELMOELPKTGKTKAVG 168
 Qy 169 SNWTLADLERK--SKPAKYMPHANOIEIHPFLPNBELYCYCFSKRNIMPVAYSPLGSQNQV 226
 Db 169 D-----SCGVREEEFVTTKLWCTQHPEPV--ALDOSLKRGLDGYDVLMLHMPARL-- 117
 Qy 227 PTTGGERVSЕНКЛНДИАЕКГНТЛЕНП-----PTVRAAMEKTYEDERKARSIGV 168
 Db 225 --DPAYTKNEDILSVPTXKDGSRAVDITNNNPKTTELMOELPKTGKTKAVG 168
 Qy 287 ADFEAINAVAK 297
 Db 283 EDFEAINTNSIK 293
 RESULT 6
 ABG93198 standard; protein; 312 AA.
 XX ABG93198;
 XX ABG93198;
 DT 21-NOV-2002 (first entry)
 XX DE *S. cerevisiae* BAX-associated protein fragment SEQ ID 354.
 XX DE S. cerevisiae BAX-associated protein fragment SEQ ID 354.
 XX DE S. cerevisiae BAX-associated protein fragment SEQ ID 354.
 XX DE S. cerevisiae BAX-associated protein fragment SEQ ID 354.

DR WPI; 1997-403701/38.
 XX Recombinant uronate reductase and L-gulono-gamma-lactone oxidase - used
 PT in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 XX
 PT Recombinant uronate reductase and L-gulono-gamma-lactone oxidase - used
 XX in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 PS Claim 31; Page; 26pp; German.

CC The present sequence represents a uronate dehydrogenase (reductase) Q9G
 CC variant. The variant can be used for synthesis of L-ascorbic acid. DNA
 CC encoding L-gulono-gamma-lactone oxidase (Gluox) is also used in the
 CC production of L-ascorbic acid. The uronate dehydrogenase has higher
 CC affinity for substrate (D-glucuronic or galacturonic acids) and coenzymes
 CC than the similar enzyme from *Lipomyces starkeyi*, also better stability
 CC and greater yield when produced recombinantly. NB. This sequence was
 CC created using the wild type uronate dehydrogenase sequence given in
 CC Figure 3 of the specification
 XX Sequence 312 AA;

Query Match 33.2%; Score 575.5; DB 2; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.5e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

Qy	5 KITFLISNGYKIPGYGFGTFASEGSKGETYATTALKTYRHLDCAWYYLNEGEVYEGFIR	64
Db	10 KILSINTAGIQIPQIGLTQSK - ENDAYKAVALTAALKDGYRIDIATAAIREDOQVEQAIIK	67
Qy	65 DFLKENPNSYKREDIFVCTKWNHLHRYEDLAVNSIDDLSIKRLGLIDYDMDFLWHPPIAEKN	124
Db	68 D----SGYPREELFVTYKLWCQHHPEV ALDQSLSKRLGLIDYDMDFLWHPPIALWMPARL--	117
Qy	125 GQGPBKIGDGKVYLKDLTENP-----PTWRAEMEKIYEDBKRASIGV	168
Db	118 -----DPAYTKNEDILSVPTKQDGSSRAVDITNNFIKTTWELQPLTGKTRAVGV	168
Qy	169 SNWTIADLERKM - SKFAKYMPHANOIEHTPFLPNEELNYQCFSKNIMPAVSPLGSONQV	226
Db	169 SNSFNLNLQDLLAQGNKLTAAQNPVQETHPFLLQDELLINFCRSKGIVVEATAPLGs---	224
Qy	227 PTTGERVSENKTNLNEIAKGNTILAQVLIAGLRGGYTLPKSSNPKRIESNFKSIELSD	286
Db	225 -TDAPLKEPVILEIAKQNVQFGHVVWYQRGYTVLPKSVNPDRIKTRNKIFTLST	282
Qy	287 ADFEAINAVAK 297	
Db	283 EDFEAINNIISK 293	

PI Matthes R, Kulbe K;
 XX WPI; 1997-403701/38.
 DR N-PSDB; AAT87004.
 XX
 PT Recombinant uronate reductase and L-gulono-gamma-lactone oxidase - used
 XX in L-ascorbic acid production from D-glucuronic or galacturonic acids.
 PS Claim 27; Page 18-19; 26pp; German.

CC The present sequence represents a uronate dehydrogenase (reductase) which is used for synthesis of L-
 CC isolated from *Saccharomyces cerevisiae*, which is used for synthesis of L-
 CC ascorbic acid. DNA encoding L-gulono-gamma-lactone oxidase (Gluox) is
 CC also used in the production of L-ascorbic acid. The uronate dehydrogenase
 CC has higher affinity for substrate (D-glucuronic or galacturonic acids)
 CC and coenzymes than the similar enzyme from *Lipomyces starkeyi*, also
 CC better stability and greater yield when produced recombinantly
 XX Sequence 312 AA;

Query Match 33.2%; Score 575.5; DB 2; Length 312;
 Best Local Similarity 40.8%; Pred. No. 1.9e-47;
 Matches 127; Conservative 54; Mismatches 85; Indels 45; Gaps 7;

Qy	5 KTFILSNGYKIPGYGFGTFASEGSKGETYATTALKTYRHLDCAWYYLNEGEVYEGFIR	64
Db	10 KILSINTAGIQIPQIGLTQSK - ENDAYKAVALTAALKDGYRIDIATAAIREDOQVEQAIIK	67
Qy	65 DFLKENPNSYKREDIFVCTKWNHLHRYEDLAVNSIDDLSIKRLGLIDYDMDFLWHPPIAEKN	124
Db	68 D----SGYPREELFVTYKLWCQHHPEV -ALDQSLSKRLGLIDYDMDFLWHPPIALWMPARL--	117
Qy	125 GQGPBKIGDGKVYLKDLTENP-----PTWRAEMEKIYEDBKRASIGV	168
Db	118 -----DPAYTKNEDILSVPTKQDGSSRAVDITNNFIKTTWELQPLTGKTRAVGV	168
Qy	169 SNWTIADLERKM - SKFAKYMPHANOIEHTPFLPNEELNYQCFSKNIMPAVSPLGSONQV	226
Db	169 SNSFNLNLQDLLAQGNKLTAAQNPVQETHPFLLQDELLINFCRSKGIVVEATAPLGs---	224
Qy	227 PTTGERVSENKTNLNEIAKGNTILAQVLIAGLRGGYTLPKSSNPKRIESNFKSIELSD	286
Db	225 -TDAPLKEPVILEIAKQNVQFGHVVWYQRGYTVLPKSVNPDRIKTRNKIFTLST	282
Qy	287 ADFEAINAVAK 297	
Db	283 EDFEAINNIISK 293	

RESULT 4
 AAW9217 standard; protein; 312 AA.
 XX AAW9217;
 AC
 XX DT 02-MAR-1998 (first entry)
 DE S. cerevisiae uronate dehydrogenase.
 KW uronate reductase; uronate dehydrogenase; *Saccharomyces cerevisiae*;
 KW synthesis; L-ascorbic acid.
 XX Saccharomyces cerevisiae.
 OS DE19504798-A1.
 XX PD 14-AUG-1997.
 XX PF 09-FEB-1996; 96DE-01004798.
 PR 09-FEB-1996; 96DE-01004798.
 XX PA (HERB-) HERBSTREITH & FOX PEKTIN-FAB NEUENBURG.

RESULT 5
 AAG63565 standard; protein; 312 AA.
 XX AAG63565;
 AC
 XX DT 15-OCT-2001 (first entry)
 DE Amino acid sequence of a nuclear gene encoding a yeast protein.
 KW Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding.
 XX Saccharomyces cerevisiae.
 OS WO200155342-A2.
 XX PN 02-AUG-2001.
 XX PD 31-JAN-2001; 2001WO-US003186.
 XX PF PR 31-JAN-2000; 2000US-0049421.
 PR 08-DEC-2000; 2000US-00734237.

Query Match	100.0%	Score 1731;	DB 5;	Length 325;				
Best Local Similarity	100.0%	Pred. No. 5.5e-161;	Mismatches 0;	Gaps 0;				
Matches 325;	Conservative 0;	Indels 0;						
Qy	1 MSNGKFTLNSGVKIPGVGFTFASERGSKGETYAVTTALKTGFRHLDCAWYTLNEGEVG 60							
Db	1 MSNGKFTLNSGVKIPGVGFTFASERGSKGETYAVTTALKTGFRHLDCAWYTLNEGEVG 60							
Qy	61 EGIRDLEKNSVKKREDIFVTKWNHLHYEDLWSIDSLSKRLGLDVTDMFLVHWT 120							
Db	61 EGIRDLEKNSVKKREDIFVTKWNHLHYEDLWSIDSLSKRLGLDVTDMFLVHWT 120							
Qy	121 AEKGCGEPKIGPDGXVYLKDLTENPEPTWAKMEKIYEDRKARSIGVSNWTADELS 180							
Db	121 AERGGCGEPKIGPDGXVYLKDLTENPEPTWAKMEKIYEDRKARSIGVSNWTADELS 180							
Qy	181 KFAKYMPHANCIEIHFPLNNEELVQCFSKRIMPVAYSPIGSQNOVPTGERYSENKTLN 240							
Db	181 KFAKYMPHANCIEIHFPLNNEELVQCFSKRIMPVAYSPIGSQNOVPTGERYSENKTLN 240							
Qy	241 EIAEKGCGNTLAQVLIWGLRGRVYVTPKSNPKIESNKSIELSDAEEAIVAKGRH 300							
Db	241 EIAEKGCGNTLAQVLIWGLRGRVYVTPKSNPKIESNKSIELSDAEEAIVAKGRH 300							
Qy	301 FREVNNKDTFSDYWPEETAKNSA 325							
Db	301 FREVNNKDTFSDYWPEETAKNSA 325							
RESULT 2								
ID	ADE39629 standard; protein; 325 AA.							
XX								
AC	ADE39629;							
XX								
DT	29-JAN-2004 (first entry)							
XX								
DE	Penicillium reducing enzyme protein, SEQ ID NO 3.							
XX								
KW	optically active 2-hydroxycycloalkane carboxylic acid ester;							
KW	enzyme; Penicillium; reduction.							
XX								
OS	Penicillium citrinum.							
XX								
PN	EPI323827-A2.							
XX								
PD	02-JTL-2003.							
XX								
PF	20-DEC-2002; 2002EP-00258814.							
XX								
PR	27-DEC-2001; 2001JP-00395884.							
PR	27-DEC-2001; 2001JP-00395885.							
PR	10-APR-2002; 2002JP-00107648.							
XX								
PA	(SUMO) SUMITOMO CHEM CO LTD.							
XX								
PI	Asako H, Wakita R, Itoh N;							
XX								
DR	WPI; 2003-723302/69.							
XX								
DR	N-PSDB; ADE39630.							
PT	Producing optically active 2-hydroxycycloalkane carboxylic acid ester by reacting 2-oxocycloalkane carboxylic acid ester with transformant having ability to asymmetrically reduce the ester to optically active ester.							
PT	Claim 9; SEQ ID NO 3; 45PP; English.							
XX								
CC	The invention relates to novel method for producing an optically active 2-hydroxycycloalkane carboxylic acid ester. The novel method involves:							
CC	allowing 2-oxocycloalkane carboxylic acid ester to react with a transformant, or a dead cell or their extract, artificially provided with							
CC	the ability to asymmetrically reduce a 2-oxocycloalkane carboxylic acid							

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

CC ester to the optically active 2-hydroxycycloalkane carboxylic acid as part

CC of the method of the invention.

Sequence 325 AA;

Score 1731; DB 5;

Length 325;

Best Local Similarity 100.0%; Pred. No. 5.5e-161;

Matches 325; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

CC ability to regenerate a coenzyme on which an enzyme having the above

CC hydroxycycloalkane carboxylic acid. The optically active 2-

CC hydroxycycloalkane carboxylic acid is useful as an intermediate for the

CC production of bioactive substances. This sequence represents a

CC Penicillium enzyme protein used to reduce 2-oxocycloalkane carboxylic acid

Result No.	Score	Query Match Length	DB ID	Description
1	1731	100.0	325 5 ABB77965	Abb77965 Protein w Ade39523 Penicilli raw29220 S. cerevi
2	1731	100.0	325 7 ADE39629	Ade39629 S. cerevi raw29217 S. cerevi
3	574.5	33.2	312 2 AAW22120	AAW22120 S. cerevi raw29220 S. cerevi
4	574.5	33.2	312 4 AAG63565	AAG63565 Amino aci abg93198 S. cerevi
5	574.5	33.2	312 5 ABG93198	ABG93198 S. cerevi
6	574.5	33.2	313 4 AAG63566	AAG63566 Synthetic abg93050 S. cerevi
7	574.5	33.2	313 4 AAG63566	AAG63566 Synthetic abg93050 S. cerevi
8	573.5	33.1	312 2 AAW22118	AAW22118 S. cerevi raw29219 S. cerevi
9	565.5	32.7	312 2 AAW22120	AAW22120 S. cerevi raw29219 S. cerevi
10	559.5	32.3	323 2 AAR96294	AAR96294 Carbonyl abg93050 T. megaloch
11	555	32.1	313 4 AAG63564	AAG63564 Amino aci abg93050 T. megaloch
12	550	31.8	295 5 ABG93030	ABG93030 C. albica abg93050 S. cerevi
13	547.5	31.6	344 5 ABG93050	ABG93050 S. cerevi
14	547.5	31.6	344 5 AAU76345	AAU76345 Yeast D-a
15	547	31.6	328 5 ABP53552	ABP53552 T. megaloch
16	545.5	31.5	328 5 ABP53551	ABP53551 T. megaloch
17	543.5	31.4	313 4 AAB47466	AAB47466 G. max al
18	543.5	31.4	313 7 AAE39122	AAE39122 Soybean a
19	541	31.3	322 4 AAG63561	AAG63561 Amino aci abg93050 T. megaloch
20	533.5	30.8	330 5 AAG07671	AAG07671 Arabidops
22	529.5	30.6	309 3 AAG33086	AAG33086 Arabidops
23	529	30.6	309 3 AAG50893	AAG50893 Arabidops
24	529	30.6	290 3 AAG31845	AAG31845 Arabidops
25	529	30.6	315 3 AAG40418	AAG40418 Arabidops

GanCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2004, 13:58:21 ; Search time 60 Seconds
 (without alignments)
 15310.466 Million cell updates/sec

Title: US-10-004-115B-1

Perfect score: 1 MSNCKTFIISNGKIPGVGR..... . MKDTEFGDVMPETAKNUSA 325

Sequence: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: GeneseqP1990:*
- 2: GeneseqP1990:*
- 3: GeneseqP2000:*
- 4: GeneseqP2001:*
- 5: GeneseqP2002:*
- 6: GeneseqP2003a:*
- 7: GeneseqP2003b:*
- 8: GeneseqP2004:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1
 ID ABB77965 standard; protein; 325 AA.
 XX
 AC ABB77965;
 XX
 DT 22-OCT-2002 (first entry)
 XX
 DE Protein which is capable of producing (S)-4-bromo-3-hydroxybutanoate.
 XX
 KW (S)-4-bromo-3-hydroxybutanoate; 4-bromo-3-oxobutanoate; Pharmaceutical; agrochemical; 4-cyano-3-hydroxybutanoic acid.
 XX
 OS Penicillium citrinum.
 XX
 PN EP1213354-A2.
 XX
 PD 12-JUN-2002.
 XX
 PP 07-DEC-2001; 2001EP-00310251.
 XX
 PR 07-DEC-2000; 2000JP-00372704.
 PR 05-JAN-2001; 2001JP-0006144.
 PR 02-FEB-2001; 2001JP-0026594.
 PR 11-JUN-2001; 2001JP-00175175.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 PI Asako H, Matsumura K, Shimizu M, Ito N, Wakita R;
 XX
 WPI; 2002-550350/59.
 DR N-PSPD; ABL59376, ABL59397, ABL59398.
 XX
 PT New protein preferentially producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-bromo-3-oxobutanoate, useful in pharmaceuticals and agrochemicals.
 PT
 XX
 BS Claim 14; Page 29-31; 56pp; English.
 XX
 CC The present sequence represents a Penicillium citrinum protein which is capable of producing (S)-4-bromo-3-hydroxybutanoate by asymmetrically reducing 4-bromo-3-oxobutanoate. The protein and poly nucleotides are useful for producing optically active (S)-4-bromo-3-hydroxybutanoate, which is useful as an intermediate in the production of pharmaceuticals and agrochemicals. The (S)-4-bromo-3-hydroxybutanoate produced can also be used to produce 4-cyano-3-hydroxybutanoic acid.
 CC
 XX Sequence 325 AA;